Communication, Society and Media
ISSN 2576-5388 (Print) ISSN 2576-5396 (Online)
Vol. 6, No. 4, 2023
www.scholink.org/ojs/index.php/csm

Original Paper

The Influence of Conversation Role on Prosodic Entrainment in Mandarin Interactions

Zhihua Xia[1*]

[1] The School of Foreign Studies, Jiangsu Normal University, Xuzhou, Jiangsu Province, China

[*] Zhihua Xia, The School of Foreign Studies, Jiangsu Normal University, Xuzhou, Jiangsu Province, China

Received: November 10, 2023 Accepted: November 22, 2023 Online Published: December 1, 2023

doi:10.22158/csm.v6n4p76 URL: http://dx.doi.org/10.22158/csm.v6n4p76

Abstract

The aim of this study is to find out how conversation role affects prosodic entrainment in Mandarin interactions. Tongji Games Corpus are adopted for this study, in which two interlocutors play unequal roles in Picture Ordering Games (information giver and information follower), and equal roles in Picture Classifying Games. Based on this corpus, two tests are conducted in this study, the Role Influence Test and Role Direction Test. In the analysis of Role Influence Test, it is found that the entrainment degree in Picture Ordering Games is significantly bigger than that in the Picture Classifying Games. In the analysis of Role Direction Test, it is found that information givers entrain more to followers in conversation. These findings provide evidences that conversation roles, as one type of the social factors, have influence on the degree and direction of prosodic entrainment in Mandarin interactions.

Keywords

prosodic entrainment, conversation role, Mandarin interaction

1. Introduction

Conversations are considered as joint activities in which two interlocutors share or synchronize their mental states and performances. The joint nature of language processing in communication requires the interpersonal coordination in minds and actions (Brennan *et al.*, 2010). This coordination is called entrainment.

Prosodic entrainment means that two speakers adapt prosody to that of their interlocutors in conversation, and then become similar in speaking for smooth and successful communication. Prosodic entrainment is also named prosodic accommodation, prosodic adaptation, or prosodic alignment

(Brennan, 1996; Cummins, 2009; Van Der Wege, 2009; Lee *et al*., 2014; Beňuš *et al.*, 2011; Levitan & Hirschberg, 2011; Levitan *et al.*, 2012; Levitan, 2013; De Looze *et al.*, 2014; Xia *et al*., 2004, 2023; Xia, 2013; Xia & Ma, 2016a, 2016b, 2019).

Few attentions in previous studies are paid to conversation roles and prosodic entrainment. Social factors are proved to have influence on prosodic entrainment (Levitan *et al*., 2013). However, most of the previous studies have focused on the gender and prosodic entrainment (Levitan *et al*., 2013, Xia *et al.*, 2014; Xia & Ma, 2016a, 2016b). Conversation role is crucial for successful communication. The previous studies have touched some types of conversational roles in other fields, for example, the sensorimotor activation related to speaker and listener role in natural conversation (Mandel *et al.*, 2016), parent-infant conversations associated with the development of preterm and term-born infants (Coughlan, 2024).

Different from the research mentioned above, this study aims to explore how the conversation role influences the prosodic entrainment in Mandarin conversation. In this study, two types of conversation roles are designed as “equal roles” in Picture Classification Games, in which two interlocutors have free discussion and make classification of all the pictures into certain categories, and the pairs take the equal roles in discussion, and “unequal roles” in Picture Ordering Games, in which two interlocutors in the conversation take the roles as information giver and information follower, and the information givers instructs followers to make correct order of the pictures.

With this research goal, the analyses of entrainment and role are accomplished by two steps. The first step is the Role Influence Test, which aims to test whether the unequal roles affect the degree of entrainment or not. If the unequal roles prove to affect the degree of entrainment in the test of first step, it is necessary to test where the differences in the degree of entrainment come from. Then the second step, called Role Direction Test, aims to test the direction of entrainment between two roles in conversations, that is to say, to find out whether givers entrain more to followers or followers entrain more to givers.

2. Variables

The present research focuses on prosody in 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).

Speaking-rate equals the number of syllables per second in this study. Syllable counts are made automatically by Praat according to the orthographic transcriptions. Provides detailed calculation of speaking-rate.

Fundamental frequency of a voice is used to measure how often the sound wave repeats itself. The present research adopts min, mean and max values of fundamental frequency, which are obtained automatically by Praat.

The features of intensity are used to describe the degree of energy in a sound wave. It is perceived as the volume of a sound. The present research adopts min, mean and max values of intensity, which are also obtained automatically by Praat.

3. Corpus

The corpus of the present research elicited by two games contains approximately 12 hours of spontaneous, task-oriented conversations between pairs of subjects. 70 pairs participated in the experiments, among which 17 pairs played the Picture Ordering game; 39 pairs played the Picture Classifying game, and 14 pairs played both. All the subjects produced 115 conversations (the pairs in Picture Ordering Game produced two conversations). The average duration of conversation is 6 minutes. The analyses of the present research were based on the conversations produced by the 70 pairs.

3.1 Unequal and Equal Roles in Conversations

Tongji Games Corpus are adopted in this study. Two games were used to elicit spontaneous conversations in this corpus. One is Picture Ordering Game, and the other is Picturing Classifying Game.

In Picture Ordering Games, two interlocutors have “unequal conversation roles”. In detail, one subject played as the information giver, and the other subject played as the information follower. The information giver’s screen displayed 18 cards in the correct order; the information follower’s screen displayed the same 18 cards but in the incorrect order. The information giver instructed the information follower to put the 18 cards in correct order. 18 pictures were labeled by the numbers. The Picture Ordering Game started from the picture numbered as 1, continued to the picture numbered as 2, 3, ... and ended by the picture numbered as 18. When the task was completed, the same pair switched roles and accomplished the picture ordering game again, in which the previous information giver played as the information follower, and the previous follower as the giver.

In Picture Classifying Game, two interlocutors have “equal conversation roles”. In detail, two interlocutors have discussion on how to classify the 18 pictures. What is presented on the screens of computers for the two interlocutors are the same (the same 18 pictures in the same order). Each pair accomplished attributive classification over 18 pictures. Interlocutors had discussion, reached agreement on the common attributes they found for the pictures, and grouped the pictures with the common attributes together. Every picture was assigned to a category by the common attributes accepted by both of the interlocutors. The classification also started from the picture numbered as 1,

continued to the picture numbered as 2, 3, ... and ended by the picture numbered as 18.

3.2 Annotation

3.2.1 IPU Segmentation

IPUs are adopted as the minimal units in present study. IPU (Inter-Pausal Unit) is the stretch of a single speaker's speech bounded by pauses (for example 100 ms). It is a minimal unit frequently used in the analyses of conversations. Koiso *et al*. (1998) utilized Inter-Pausal Units (IPUs) as the minimal units in the research of intonation in the turn taking in Japanese. Casper (2003) also employed this unit in studying local speech melody in turn taking system in Dutch. Levitan & Hirschberg (2011), Levitan *et al.* (2012) annotated the conversations by IPUs (pause-free units of speech from a single speaker separated from one another by at least 50 ms) in the study of prosodic entrainment in English.

The threshold for IPUs of the present research is 80 ms. About setting the threshold of IPUs, the previous studies point out that the threshold can't be too big, because the longer pause in IPUs would cause the omission of the IPUs with shorter pauses, and the prosodic descriptions would be ignored; the threshold can't be too small, because it would be difficult to distinguish the real pause and the stop gap within segments, including plosives, affricatives, etc. So, the threshold should be between these two extremes. It should be as small as possible, but it should be longer than the maximal stop gap of segments. Therefore, the threshold of IPUs is the maximal stop gap of segments in the corpus. Thus, to find the threshold of IPUs is to find the maximal stop gap of plosives at the word initial position in Chinese conversations of this corpus. This method was carried out in several steps in the present research. The first step was to choose at random the recordings from 6 female and 6 male subjects. The second step was to find the stop gaps of plosives including [p^h], [p], [t^h], [t], [k^h], [k] at the word initial positions in these 12 recordings. The third step was to measure the durations of these stop gaps. The fourth step was to calculate the maximal stop gap. The maximum is 74.2 ms, and the present research set the closest integer value as 80 ms.

IPU segmentation of the present study is accomplished in two steps. The first step is automatic labelling by SPPAS (Bigi & Hirst, 2012). The second step is the check of IPUs' boundaries in Praat.

3.2.2 Annotation in Praat

The annotation work in Praat of the present research includes 4 steps.

The first step is a check on IPU boundaries.

The second step is dealing with the spontaneous speech phenomena in conversation.

The third step is dealing with the boundaries of initial plosives. After all these three steps, the boundaries of IPUs are fixed.

Then, the last step is to input Mandarin characters within IPUs for the data extraction in the following task.

Data extraction was operated on these 7 variables over every IPU by Praat script.

4. Results and Discussions

4.1 Role Influence Tests

This analysis is to find out: Does the unequal roles affect the degree of entrainment?

The unequal roles in this section are restricted to the information giver and information follower. Entrainment in conversation is proved to be related closely to social factors, and the individual differences of the interlocutors decide the direction and state of entrainment in conversation. Therefore, it is hypothesized that the unequal roles affect the degree of entrainment in conversation.

4.1.1 Paired T-tests between *odis* and *cdis*

In order to find out whether the unequal roles affect the degree of entrainment or not, the present research compares the entrainment degree of the pairs with unequal roles and that of the pairs with equal roles, which are controlled in Picture Ordering Games and Picture Classifying Games. So, research question in this section is converted to find whether the entrainment degree is different between Picture Ordering Games and Picture Classifying Games or not.

It is hypothesized that the unequal roles affect the degree of entrainment.

In order to analyze the role influence in the Picture Ordering Games and Picture Classifying Games, paired T-test is used in comparing entrainment degree between these two games. As mentioned above, the present research has made the control that the same pairs play both games, so the interference from the different pairs of speakers is controlled for in this analysis. Thus, the paired T-tests over the same interlocutors are valid to find out the difference of entrainment degree between two games.

The entrainment degree in Picture Ordering Games, *odis*, refers to the distance of prosodic features over the whole conversation between two speakers in Picture Ordering Games. The entrainment degree in Picture Classifying Games, *cdis*, refers to the distance of prosodic features over the whole conversation between two speakers in Picture Classifying Games. Paired T-tests are accomplished between *odis* and *cdis*.

According to the description of corpus, pairs of subjects participated in both the Picture Ordering Games and Picture Classifying Games.

In terms of *odis*, the distance of prosodic features over the whole conversation between two speakers in the Picture Ordering Games conversations, the average distance from two conversations is calculated as its value, because every pair of subjects plays the Picture Ordering Game twice by switching the roles from information giver to information follower in Picture Ordering Games. Thus, 14 *odis* are computed for one prosodic parameter over 28 conversations in Picture Ordering Games by getting their averages.

In terms of *cdis*, the distance of prosodic features over the whole conversation between two speakers in Picture Classifying Games conversations, one value of *cdis* comes from one conversation, because every pair plays classifying game once. 14 pairs participated in 14 conversations, so 14 *cdis* are computed over 14 conversations for one prosodic parameter.

One paired T-test is accomplished between 14 *odis* and 14 *cdis* for one prosodic parameter. 7 paired T-tests are accomplished for 7 prosodic parameters.

The data used in paired T-tests are listed in the following Table 1 and 2.

Table 1. The Values of *odis* in Picture Ordering Games

	Speaking _rate	F0_min	F0_mean	F0_max	Int_min	Int_mean	Int_max
1	1.36	24.56	19.43	31.92	6.54	7.70	7.83
2	0.33	19.38	31.72	32.16	14.02	9.78	9.07
3	0.10	13.06	8.50	15.25	1.55	7.16	8.84
4	1.09	12.82	26.33	56.78	5.17	3.85	3.83
5	1.09	78.38	104.50	152.71	8.60	0.59	0.19
6	0.59	27.76	45.63	57.65	2.64	7.31	8.26
7	0.98	29.28	28.73	37.18	3.24	4.48	4.34
8	1.18	21.78	21.20	22.51	6.79	4.19	4.82
9	0.19	13.80	14.38	21.62	2.39	2.95	3.16
10	1.09	2.97	13.71	32.23	3.69	4.01	4.31
11	0.70	6.08	1.64	5.93	2.04	3.71	3.73
12	0.87	7.23	6.14	15.05	1.48	1.47	1.26
13	0.85	26.81	25.72	23.99	2.88	6.79	7.13
14	0.46	16.57	14.00	30.11	2.68	2.63	3.24

Table 2. The Values of *cdis* in Picture Classifying Games

	Speaking _rate	F0_min	F0_mean	F0_max	Int_min	Int_mean	Int_max
1	1.66	19.49	25.55	38.02	11.45	6.05	5.64
2	0.30	11.65	27.60	33.57	11.92	11.02	10.45
3	0.34	0.71	1.89	4.12	3.74	3.08	3.80
4	0.51	5.08	18.68	44.65	9.07	4.78	3.62
5	1.26	72.52	105.69	150.37	13.43	1.55	1.77
6	0.15	26.78	49.28	64.88	3.03	8.79	9.70
7	0.29	21.09	22.59	28.09	0.54	2.65	3.33
8	0.55	15.70	18.67	21.59	7.14	2.19	1.89
9	0.78	15.13	13.22	20.45	1.26	2.18	2.21
10	0.54	4.54	10.07	28.00	4.31	3.31	3.71

11	0.15	1.01	0.95	1.70	1.95	1.32	1.55
12	0.40	11.54	6.96	6.47	3.25	2.87	3.37
13	0.05	24.31	20.15	12.65	5.12	9.19	9.13
14	0.29	18.81	6.40	21.38	0.56	2.46	2.97

4.1.2 Results of the Entrainment Degree Tests

The results in paired T-tests of entrainment degree for the same interlocutors between two games are listed in Table 3.

Table 3. Paired T-tests in Entrainment Degree between Two Games

Feature	*t*	*df*	*p-value*	*Sig.*	*meano*	*meanc*
Speaking-rate	2.225	13	0.044	*	0.7782	0.5194
F0 min	2.887	13	0.013	*	21.4646	17.7399
F0 mean	2.119	13	0.054	/	25.8304	23.4078
F0 max	2.506	13	0.026	*	38.2213	33.9959
Intensity min	-1.391	13	0.187	/	4.5517	5.4831
Intensity mean	0.746	13	0.469	/	4.7587	4.3888
Intensity max	0.875	13	0.397	/	5.0007	4.5094

In Table 3, an asterisk * indicates the significant difference, and the symbol / indicates no significant difference. According to this table, the entrainment degree between the Picture Ordering Games and Picture Classifying Games are significantly different over 3 prosodic parameters: Speaking-rate ($p= 0.044 < 0.05$), F0 min ($p= 0.013 < 0.05$) and F0 max ($p= 0.026 < 0.05$). The entrainment degree between two games is significantly different over F0 mean ($p= 0.054 < 0.1$) at confidence level 0.1.

These results indicate that the degree of entrainment is significantly different between Picture Ordering Games and Picture Classifying Games over three prosodic parameters: Speaking-rate, F0min, and F0max, and the feature of F0 mean tends to be different between the two games.

In addition, the mean entrainment degree in picture ordering game (indicated by *meano* in Table 2) is bigger than the mean entrainment degree in picture classifying game (indicated by *meanc* in Table 3) over almost every prosodic feature tested over the feature of duration and F0. It is the same for the features of intensity, although the differences are not significant.

These indicate that the entrainment degree in the picture ordering game is larger than the entrainment degree in the picture classifying game over almost every prosodic feature tested.

The results above prove that the unequal roles (information giver or follower) affect the entrainment degree in conversation over the prosodic features of duration and F0, and the entrainment in the picture ordering games is stronger than that in the picture classifying games. That is, there is more entrainment in the conversations with unequal roles than in those with equal roles.

4.2 Role Direction Tests

It is proved in the Section 3.1 that the unequal roles (information giver or follower) affect the entrainment degree in conversation over the prosodic features of duration and F0. Questions can be asked further: what is the entrainment direction? Do information givers entrain more to followers, or do followers entrain more to givers?

4.2.1 Paired T-tests between ｜Ag-A0｜and ｜Af-A0｜

As mentioned at the beginning of this section, the Role Direction Test is the second step in the analyses of entrainment and role. In the first step, the Role Influence Test, it is found that the degree of entrainment is significantly different between Picture Ordering Games and Picture Classifying Games over the features of duration and F0. Then it is necessary to analyze where the differences in the degree of entrainment come from.

As mentioned above, some subjects are chosen to participate in both Picture Ordering Games and Picture Classifying Games. In Picture Ordering Games, two interlocutors (A and B) have different roles as information giver and information follower. When they finished the first game, they switched the roles in another ordering game. That is, two interlocutors participated in two ordering games, in which they switched the roles between information giver and information follower. In picture classifying, two speakers have the same roles in discussion. Thus, every subject has 3 speaking states.

For instance, a subject in a pair (A & B) has 3 speaking states.

Ag: A plays as an information giver in the first picture ordering game with B.

Af: A plays as an information follower in the second picture ordering game with B.

A0: A plays in an equal role with B in a picture classifying game.

Bg: B plays as an information follower in the first picture ordering game with A.

Bf: B plays as an information giver in the second picture ordering game with A.

B0: B plays in an equal role with A in a picture classifying game.

The same pair plays two kinds of games: the Picture Classifying Games and Picture Ordering Games. Taking one side of the interlocutors as an example, A0 is A’s speaking state with equal roles; Ag is A’s speaking state when A is an information giver; Af is A’s speaking state when A is an information follower. The difference between Ag and Af is the change in speaking state caused by the change in A’s roles from the information giver to the information follower, while the partner is still B. What is changed in B in Picture Ordering Games is also the state of speaking from the information follower to

the information giver. Therefore, as the other side of the pair, the partner of A, B also has 3 similar speaking states.

The differences in the entrainment degree between two kinds of games found in the Section 5.1 are assumed to be caused by the difference in the roles. In the present research, among the 3 speaking states of A, A0 is considered as a reference state, because this speaking state is produced under the equal role with the partner. The deviation between Ag and A0 is caused only by A's role of a giver, and the deviation between Af and A0 is caused only by A's role of a follower. Therefore, | Af-A0 | is the deviation caused by the role of giver; | Af-A0 | is the deviation caused by the role of follower. The key point to make such comparisons valid is that in this analysis the partner of A is not changed, and what is changed is only the role of A. And all of this attributes to the experiment design of the present research that several pairs of subjects participated in both games.

A paired T-test between | Ag-A0 | and | Af-A0 | is adopted in this section in order to find out which role causes bigger deviation. If | Ag-A0 | is bigger than | Af-A0 |, the deviation caused by the role of giver is bigger than that caused by the role of follower. That is, the giver makes larger deviation in entrainment than the follower does in conversation, so the giver entrains more to the follower. If | Ag-A0 | is smaller than /Af-A0/, the deviation caused by the role of giver is smaller than that caused by the role of follower. That is, the follower makes a larger derivation in entrainment than the giver does in conversation, so the follower entrains more to the giver.

One point should be mentioned in the use of this method. The post-conversation influence is not considered in this analysis. Fourteen pairs participated in both the Picture Ordering Games and Picture Classifying Games. Every two interlocutors of them participated in two ordering games, in which they switched the roles between information giver and information follower, and a picture classifying game, in which two speakers in the pair have the same roles in discussion. One pair of these subjects actually produced three conversations. Thus, a question is raised: is there post-conversation influence among these conversations? That is, is the entrainment state in the followed conversations influenced by the previous ones?

In the present research, the post-conversation influence of these three conversations produced by every pair of the 14 is not considered, because the requirements of these games are different. After the first picture ordering game, the roles of two interlocutors are changed in the second picture ordering game, and the content of picture classifying game is quite different from the previous two. Thus, the requirements of the games are always different. Therefore, the post-conversation influence is excluded in these analyses.

According to the corpus description in Section 2, 14 pairs (28) subjects participated in both the Picture Ordering Games and Picture Classifying Games. Then 28 subjects made 28 conversations. One subject (A) has 3 speaking states, Ag, Af and A0 from the Picture Ordering Games and Picture Classifying

Games. For 28 subjects, there are 28 groups of Ag, Af and A0.

The role influence test in Section 3.1 is the precondition of the Role Direction Test in Section 3.2. That is, those features which show role influence could be tested further for the role direction. According to the results of role influence test, different entrainment degree is found between unequal roles (the role of information giver and follower) and equal roles (the role of picture classifier) over the features of duration and F0. So as the further test of Role Influence Test, in the Role Direction Test, paired T-tests are accomplished over the features of duration and F0 including Speaking-rate, F0min, F0mean, and F0max.

The values of ｜Ag-A0｜and ｜Af-A0｜used in paired T-test are listed in the following Table 4.

Table 4. The values of ｜Ag-A0｜and ｜Af-A0｜

	｜Ag-A0｜				｜Af-A0｜			
	Speaking_rate	F0_min	F0_mean	F0_max	Speaking_rate	F0_min	F0_mean	F0_max
1	0.34	5.40	1.01	11.92	0.18	19.03	10.64	4.00
2	0.98	1.93	20.62	33.59	0.12	1.55	1.23	5.46
3	0.36	2.90	3.30	1.31	0.14	6.42	1.86	9.68
4	0.16	5.27	0.88	8.18	0.00	0.87	5.91	13.74
5	0.29	2.90	2.62	8.67	0.07	10.90	3.06	3.26
6	0.08	5.66	12.74	26.45	0.07	6.66	6.16	1.36
7	0.23	7.66	17.63	27.87	0.60	22.36	4.84	6.97
8	0.24	4.46	1.84	2.52	0.09	15.40	5.32	0.83
9	1.16	2.91	0.93	0.32	0.13	8.47	5.85	2.23
10	0.49	7.20	2.80	18.33	0.67	10.08	0.27	11.74
11	0.72	1.77	1.69	2.82	0.22	3.96	0.89	0.02
12	0.26	5.85	9.95	16.97	0.20	1.70	18.06	34.23
13	1.30	3.31	3.68	6.80	0.20	9.66	0.69	13.59
14	1.56	2.24	3.79	11.89	0.70	1.18	5.50	13.07
15	0.68	1.94	4.37	4.95	0.02	0.24	1.01	1.48
16	0.83	6.23	8.45	10.47	0.87	7.62	2.00	2.20
17	0.03	1.91	7.56	8.39	0.67	6.50	2.47	2.95
18	1.35	7.36	5.48	4.95	1.03	3.73	2.24	1.83
19	0.81	5.37	7.03	11.39	0.13	8.50	3.41	2.02
20	0.09	1.42	3.53	6.37	0.07	0.26	0.37	1.42
21	0.03	3.20	0.51	1.34	0.46	3.68	0.83	2.82

22	0.57	3.25	0.84	4.99	0.33	2.04	1.90	0.71
23	0.75	5.57	14.72	24.58	0.83	6.81	8.23	5.04
24	0.41	2.22	12.23	25.72	0.47	1.53	9.07	21.07
25	0.26	1.53	5.77	10.86	0.20	5.38	4.19	3.33
26	1.20	1.72	1.44	0.94	0.86	0.18	2.62	9.43
27	0.04	1.98	7.46	26.30	0.03	0.21	2.66	2.19
28	0.09	0.47	13.40	35.19	0.31	2.23	4.48	3.46

4.2.2 Results of the Role Direction Tests

In Table 4, for the features of duration and F0, 28 groups of ｜Ag-A0｜and ｜Af-A0｜are listed. The paired T-tests are accomplished between these groups.

The results of paired T-tests between ｜Ag-A0｜and ｜Af-A0｜are listed in Table 5.

Table 5. Paired T-tests between ｜Ag-A0｜and /Af-A0/

Feature	*t*	*df*	*p-value*	*Sig.*	｜Ag-A0｜	｜Af-A0｜
Speaking-rate	2.481	27	0.02	*	0.5472	0.3448
F0 min	-2.437	27	0.8	/	3.7020	5.9698
F0 mean	1.912	27	0.067	/	6.2957	4.1347
F0 max	2.790	27	0.01	*	12.6457	6.4339

In Table 5, an asterisk * indicates the significant difference, and the symbol / indicates no significant difference. According to this table, the ｜Ag-A0｜is significantly and consistently bigger than /Af-A0/ over speaking rate and F0 max ($p=0.02$, $0.01 < 0.05$). In addition, ｜Ag-A0｜is significantly bigger than ｜Af-A0｜over F0 max ($p=0.067 < 0.1$) at confidence level 0.1. These show that Ag makes the significantly bigger deviation than Af. That is, when A plays as the information giver, deviation in entrainment is bigger than that when A plays as the information follower.

The results in Table 4 indicate that givers make larger adjustments towards their followers, and followers make smaller adjustments towards their givers. When a giver makes larger adjustment and a follower makes smaller adjustment, the giver entrains more to the follower.

Based on the results above, it is found that there is more entrainment when speakers are givers than when they are followers. Information givers entrain more than followers in conversation.

In conversation, givers entrain more to followers. And the entrainment from givers to followers is prominent over Speaking-rate and F0 max which are the main features that show global entrainment as well.

5. Conclusion and Discussions

5.1 Major Findings

This study has accomplished two tests: Role Influence Test and Role Direction Test.

Through the Role Influence Test, it is found that the unequal roles (information giver or follower) affect the entrainment degree in conversation over the prosodic features of duration and F0, and there is more entrainment in the conversations with unequal roles than in those with equal roles.

Through the Role Direction Test, it is found that there is more entrainment when speakers are givers than when they are followers. Information givers entrain more than followers in conversation.

In conversations, givers entrain more to followers. And the entrainment from givers to followers is prominent in Speaking-rate and F0 max.

Through the Role Influence Test and Role Direction Test, this study has the following major findings for conversation roles and prosodic entrainment:

1) The unequal roles (information giver or follower) affect the entrainment degree in conversation over the prosodic features of duration and F0, and there is more entrainment in the conversations with unequal roles than in those with equal roles.

2) There is more entrainment when speakers are givers than when they are followers. Information givers entrain more than followers in conversation. In conversation, givers entrain more to followers. And the entrainment from givers to followers is prominent in Speaking-rate and F0 max.

5.2 Discussions

These results supply evidence for the fact that there is directive distinction in entrainment between two interlocutors. That is, the entrainment degree of a pair in conversation is unbalanced. We can feel the entrainment direction in our daily life. For instance, in Child-directed speech or motherese (Fernald *et al*., 1989), which is a term used to describe speakers' accommodation when having conversations with children. It is observed that adults entrain more to children in their conversations. For another example, in foreign talk or foreigness (Ferguson, 1975; Zuengler, 1991; Smith, 2007), which is used when talking with non-native speakers, we observe that native speakers entrain more to the foreigners.

The analyses of prosodic entrainment and role (information givers and information followers) in the present research support the existence of directive distinction in entrainment between two interlocutors as the previous studies in literature. Motivated by the hypothesis of *audience design* (Bell, 1984), which means the audience-oriented entrainment, studies have been accomplished to find out to how and what extent speakers adapt to their addressees (Branigan *et al*., 2007; Galati & Brenna, 2010). Kristiansen (2008) proposes that there are three possible directions in entrainment between two interlocutors (A, B) in conversation. One is A to B (A→B), the other one is B to A (A←B), and the third one is that both A and B approach to an intermediate state (A→←B).

There exists unbalanced entrainment between two interlocutors in conversation. It is necessary to do research on the production and perception of the unbalanced entrainment in conversation.

Different from the analyses in previous studies, the present research has examined the entrainment between information givers and information followers. In this chapter, Role influence test is made to examine whether there is prosodic entrainment between information givers and information followers, based on which, Role direction test is made to find out the direction of entrainment between them, that is, whether information givers entrain to information followers, or followers entrain to givers. It is found by the present research that more entrainment is made from givers to followers.

The present research analyzes the role in prosodic entrainment. These speakers' roles are controlled in this study. 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 conversation role in this study and gender in previous studies, more social factors should be involved in the future study of prosodic entrainment.

The emphasis and difficulty in research on the various social factors in entrainment is how to appropriately control these factors in experiments. In order to make some factors prominent, it is necessary to control for other factors. It is usually not easy to control several social factors in one experiment. Efforts should be made to produce well balanced experimental design in order to control some social factors in one experiment or a series of experiments. Much future work could be done in this direction.

It is necessary to have discussion on the methods used in this study. What should be mentioned is the method used in Role Direction Test.

In this test, a subject in a pair (A & B) has 3 speaking states (A0, Ag, and Af). Taking one side of the interlocutors as an example, A0 is A's speaking state with equal roles; Ag is A's speaking state when A is an information giver; Af is A's speaking state when A is an information follower. The difference between Ag and Af is the change in speaking state caused by the change in A's roles from the information giver to the information follower.

Because several pairs of the subjects are asked to play two kinds of games, the interference from the speakers is reduced. In the present research, A0 is considered as a reference state, because this speaking state is produced under the equal role with the partner. The deviation between Ag and A0 is caused only by A's role of a giver, and the deviation between Af and A0 is caused only by A's role of a follower. Therefore, /Ag-A0/ is the deviation caused by the role of giver; /Af-A0/ is the deviation caused by the role of follower. The key point to make such comparisons valid is that in this analysis the

partner of A is not changed, and what is changed is only the role of A. And all of this attributes to the experiment design of the present research that several pairs of subjects participated in both games.

However, there are limitations in the use of this method.

Firstly, the number of subjects who participated in both games should be enlarged. In the present research, 14 pairs of subjects participated in both the Picture Ordering Games and Picture Classifying Games. Then 28 subjects made 28 conversations. One subject (A) has 3 speaking states, Ag, Af and A0 in the Picture Ordering Games and Picture Classifying Games. For 28 subjects, there are 28 groups of Ag, Af and A0. Therefore, there 28 groups of | Af-A0 | and | Af-A0 | in the paired T-tests. 28 values are not the big samples in statistics. The number should be larger. That is, there should be more subjects who participate in both games in the future experiments.

Secondly, quantitative analyses are needed to test post-conversation influence between the conversations produced by each pair in this analysis. As mentioned in the Section 3.2.1, the post-conversation influence of the three conversations produced by every pair of the 14 is not considered in this analysis, because the requirements of these games are different. In detail, after the first picture ordering game, the roles of two interlocutors are changed in the second picture ordering game, and the content of picture classifying game is quite different from the previous two. However, we can't deny the existence of the post-conversation influence between conversations. Quantitative analyses are needed to test it in the future studies.

5.3 Conclusions

Conversation roles are proved to have influence on prosodic entrainment in Mandarin conversations in this study. More comprehensive tests are suggested to examine the influence of conversation roles on prosodic entrainment, and other social factors are also needed to be tested in the study of pragmatic use of social factors and prosodic entrainment. At the same time, the cross-linguistic comparisons of the conversation role's influence on prosodic entrainment are also expected in the future. The findings of all these studies are of crucial importance for exploring the nature of human conversations and for supplying references to construction better models of human-machine interactions.

Acknowledgments

This work was supported by the National Social Science Fund of China (NSSF Grant No. 20BYY099), the Fund for English Teaching from Jiangsu Higher Education Association (2022WJZD006), the Fund for Courses Construction in Ideological and Political Education from Jiangsu Normal University (22KCSZkc10), and the Fund for English Lesson Construction from the School of Foreign Studies in Jiangsu Normal University (2022YYKC3, 2022YYKC6).

References

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*(2), 145-204.

Beňuš, Š., Gravano, A., & Hirschberg, J. (2011). Pragmatic aspects of temporal accommodation in turn-taking. *Journal of Pragmatics*, *43*, 3001-3027.

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.

Branigan, H. P., Pickering, M. J., McLean, J. F., & Cleland, A. A. (2007). Syntactic alignment and participant role in dialogue. *Cognition*, *104*(2), 163-197.

Brennan, S. E. (1996). Lexical entrainment in spontaneous dialog. In *Proceedings of International Symposium on Spoken Dialogue* (*ISSD* pp.41-44). Philadelphia.

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two Minds, One Dialog: coordinating Speaking and Understanding. In H. R. Brian (Ed.), *The Psychology of Learning and Motivation* (Vol. 53, pp. 301-344). Academic Press.

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

Coughlan, S., Quigley, J., & Nixon, E. (2024). Parent-infant conversations are differentially associated with the development of preterm and term-born infants. *Journal of Experimental Child Psychology*, *239*, 105809.

Cummins, F. (2009). Rhythm as entrainment: The case of synchronous speech. *Journal of Phonetics*, *37*, 16-28.

De Looze, C., Scherer, S., Vaughan, B., & Campbell, N. (2014). Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction. *Speech Communication*, *58*, 11-34.

Ferguson, C. A. (1975). Toward a characterization of English foreigner talk. *Anthropological Linguistics*, *17*(1), 1-14.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., de Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(03), 477-501.

Galati, A., & Brenna, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*, 35-51.

Koiso, H., Horinchi, Y., Tutiya, S., Ichikawa, A., & Den, Y. (1998). An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japaneses Map Task dialogues. *Language and Speech*, *41*, 295-321.

Kristiansen, G. (2008). Style-shifting and shifting styles: A socio-cognitive approach to lectal variation. In D. Geeraerts, R. Dirven, & J. R. Taylor (Eds), *Cognitive Sociolinguistics* (pp.45-88). Berlin: Mouton de Gruyter.

Lee, C. C., Katsamanis, A., Black, M. P., Baucom, B. R., Christensen, A., Georgiou, P. G., & Narayanan, S. S. (2014). Computing vocal entrainment: A signal-derived PCA-based quantification scheme with application to affect analysis in married couple interactions. *Computer Speech & Language*, *28*(2), 518-539.

Levitan, R. (2013). Entrainment in spoken dialogue systems: Adopting, perdicting, and influencing user behavior. In *Proceedings of the NAACL HLT 2013 Student Research Workshop* (pp. 84-90), Atlanta, Georgia.

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.

Levitan, R., Gravano, A., Willson, L., Benus, S., 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.

Smith, C. (2007). Prosodic accommodation by French speakers to a non-native interlocutor. In *Proceedings of the XVIth International Congress of Phonetic Sciences* (pp.313-348).

Mandel, A., Bourguignon, M., Parkkonen, L. & Hari, R. (2016). Sensorimotor activation related to speaker and listener role during natural conversation. *Neuroscience Letters*, *614*, 99-104.

Van Der Wege, M. M. (2009). Lexical entrainment and lexical differentiation in reference phrase choice. *Journal of Memory and Language*, *60*, 448-463.

Xia, Z. H. (2013). Prosodic realization of topic structure in Chinese EFL learners' speaking—An empirical research based on spontaneous speeches. *Foreign Language Teaching and Research*, *45*(3), 398-410.

Xia, Z. H., Hirschberg, J., & Levitan, R. (2023). Investigating prosodic entrainment from global conversations to local turns and tones in Mandarin conversations. *Speech Communication*, *153*.

Xia, Z. H., Levitan, R., & Hirschberg, J. (2014). Prosodic entrainment in Mandarin and English: A cross-linguistic comparison. In *Proceedings of Speech Prosody* (pp. 65-69). Dublin, Ireland.

Xia, Z. H., & Ma, Q. W. (2019). *Prosodic Entrainment in Mandarin Chinese Conversations: An Experimental Study*. Tongji University Press, Shanghai, China.

Xia, Z. H., & Ma, Q. W. (2016a). The influence of gender on prosodic entrainment in Mandarin conversations. *Journal of Phonetics in China*, 118-128.

Xia, Z. H., & Ma, Q. W. (2016b). *Gender and prosodic entrainment in Mandarin conversations*. In *Proceedings of 2016 10th International Symposium on Chinese Spoken Language Processing*

(ISCSLP 2016). Tianjin, China.

Zuengler, J. (1991). Accommodation in native-non-native interactions: Going beyond the "what" to the "why" in second language research. In H. Giles, J. Coupland, & N. Coupland (Eds.), *Contexts of Accommodations* (pp. 223-244). Cambridge: Cambridge Press.