

Prosodic Entrainment in Mandarin Chinese Conversations: An Experimental Study

XIA Zhihua

Abstract This research is to explore prosodic entrainment in Mandarin Chinese conversations. Based on Tongji Games Corpus, three main analyses of prosodic entrainment are accomplished, which include prosodic entrainment at multiple levels, prosodic entrainment and social factors, and cross-linguistic comparison of prosodic entrainment. Major findings are listed as follows: 1) Two kinds of entrainment are proposed: absolute entrainment and relative entrainment; 2) Entrainment in Mandarin conversation is proved to exist in 3 main aspects of prosody; 3) Different prosodic features exhibit entrainment at different levels; 4) For the prosodic features measured in the present research, local entrainment is more evident than global entrainment; 5) With the progress of conversation, increased similarity between two interlocutors is found at turn level, but no increased similarity is found at conversation level or over tone units; 6) Entrainment is proved to be closely related to social factors; 7) Striking similarities in entrainment patterns are found in Mandarin Chinese and Standard American English, and differences in entrainment degree are found in both languages.

Key Words prosody, entrainment, interaction, pragmatic function, cross-linguistic comparison

1. INTRODUCTION

Having conversation is a joint action in which interacting individuals coordinate their behavior and adapt their linguistic choices to each other. Often this entrainment or accommodation produces convergence in conception, syntactic forms, lexicon choices, prosody, pronunciations, postures and other behavior of interlocutors. The present research focuses on prosody in 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.

Prosodic entrainment is ubiquitous in social interaction because social interaction is a joint and dynamic activity in which all participants coordinate their behavior in order to establish mutual understanding. The process

of social interaction is inherently complicated, because it involves cognitive, psychological, linguistic and sociological factors, which allow interlocutors to produce and perceive their intension, accomplish mutual understanding and establish a social relation. As an essential component in speaking, prosody plays a critical role in social interaction.

The main goal of the present research is to explore prosodic entrainment in Mandarin Chinese conversations. With this research goal, three research questions are put forward: How is prosody used to entrain in Mandarin conversation? How do social factors affect prosodic entrainment in Mandarin conversation? How is prosodic entrainment in Mandarin conversation differ or similar to that in English conversation? Therefore, three main analyses of prosodic entrainment are accomplished in the present study, which include prosodic entrainment at multiple levels, prosodic entrainment and social factors, and cross-linguistic comparison of prosodic entrainment.

2. SIGNIFICANCE

Significance of this research exists in both theoretical and practical sides.

Theoretical significance of the present research comes from two aspects.

Firstly, the study of prosodic entrainment in Mandarin conversation verifies and enriches the theoretical construction of entrainment. Chinese has different tone and prosodic systems from other types of languages. The way that prosody works in Mandarin Chinese verifies theories of entrainment based on the previous researches. In addition, classification of absolute and relative entrainment that firstly proposed in the present research enriches the theoretical construction of entrainment. Both absolute and relative entrainment is proved in the prosody of Mandarin Chinese conversations in the present research, in which absolute entrainment is found over global units, and relative entrainment between the relative pitch registers of tones in the conversations of mixed gender pairs. The proposal of two-category entrainment in collaborative communication is beneficial to description and understanding of entrainment process in conversation.

Secondly, the cross-linguistic comparisons of prosodic entrainment between Chinese and English are of significance in linguistic typology. From typological viewpoint, Chinese and English are different. It is worthy to do comparison of prosodic entrainment between the two in order to explore whether they have similarities or differences on pragmatic functions of prosody in different types of language. Thus, the present research helps people make deeper understanding of prosodic entrainment in different languages.

Practical significance of the present research comes from two aspects.

At first, strategies and instructions for successful communication can be found based on the results of this research. The prosodic entrainment involves speakers' mutual cooperation and coordination. People can take the prosodic parameters including the features of duration, F0 and intensity, which are analyzed in this study, as effective coordination means,

and make appropriate adjustment in communication.

In addition, the results of this study are expected to bring certain benefits in speech synthesis. Making synthesized language as natural as that of human beings is one of the challenges in speech synthesis. Fine-grained adjustment of prosody in entrainment found in this research is of practical significance to improve the synthesized language in communication.

3. CORPUS

The corpus and the main analyses of the present research are listed as follows:

The present research is based on Tongji Games Corpus. The corpus contains 115 spontaneous, task-oriented conversations between pairs of subjects, average duration of which is 6 minutes. Experiments are designed to elicit spontaneous and collaborative conversations. Subjects accomplish two kinds of games (Picture Ordering and Picture Classifying), which stimulate the cooperative conversations. 70 pairs participated in the experiments.

The present research sets 7 acoustic-prosodic features from the 3 main aspects of prosody as the parameters in analyses: Speaking-rate from durational features; F0 min, F0 mean, and F0 max from the feature of F0; Intensity min, Intensity mean, and Intensity max from the feature of intensity.

IPU (the minimal pause interval is set as 80 ms in the present study) is identified as the smallest analysis unit in this research. After the automatic annotation of IPUs by SPPAS, the boundaries of IPUs are checked by professional annotators. The values of seven prosodic features in every IPU are extracted by Praat. These values constitute the basic data for analyses. When the units in analyses contain more than one IPU, the weighted averages of IPUs within these units are calculated (weights are durations of IPUs). Then, the weighted averages are used as the values in analyses.

4. ANALYSES

The analyses of entrainment at multiple levels are presented first. From the global to the local, the analyses of prosodic entrainment in this part are accomplished at three levels (the levels of conversations, turns, and tone units). The analyses at each level are made from 3 main aspects including proximity, convergence, and synchrony. In addition, cross-level comparisons of the results from the same aspects at different levels are made.

Through the comparison from the aspect of proximity, it is found that the features of duration and intensity show proximity at all the levels. Through the comparison from the aspect of convergence, it is found that (i) it is more difficult to capture convergence in conversation than proximity and synchrony; (ii) convergence is found at turn level, and features of F0 show better prosodic convergence than the feature of duration and intensity. Through the comparison from the aspect of synchrony, it is found that synchrony is possible even without proximity and convergence. That is, interlocutors entrain over local units on certain features, irrespective of their global similarity.

The analysis of entrainment over tones proves relative entrainment between the relative pitch registers of tones in the conversations of mixed gender pairs.

The analyses of relationship between entrainment and social factors include two analyses: the analysis of entrainment and gender, and the analysis of entrainment and role.

The analyses of entrainment and gender include two parts: proximity by gender groups, and entrainment degree by gender groups. In the analysis of proximity by gender groups, it is found the most prosodic features are entrained in female-male conversations, and the least in male-male conversations; in the same gender pairs' conversations (female-female and male-male conversations), the features of duration and intensity are entrained; in mixed gender pairs' conversations (female-male), besides the same features entrained above, features of F0 are also entrained. In the analysis of entrainment degree by gender

groups, it is found that male-male pairs tend to entrain least in MC conversations.

The analyses of entrainment and role include two parts: 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. Two interlocutors play unequal roles in Picture Ordering Games (information giver and information follower), and pairs play equal roles in Picture Classifying Games. This result of Role Influence Test shows that different roles have influence on the entrainment degree. In the analysis of Role Direction Test, it is found that information givers entrain more to followers in conversation.

Cross-linguistic comparison of prosodic entrainment between Mandarin Chinese and Standard American English is presented at last. Based on the parallel work in these two languages, the comparisons of prosodic entrainment are made at both global and local levels (the levels of conversation and turn) and over different gender groups.

Through the cross-linguistic comparison, striking similarities are found in entrainment patterns in both languages. For example, at conversation level, there are similar proximity patterns in conversation of MC and SAE; at turn level, in both languages, holistic proximity between interlocutors over turns is found over the same prosodic features including Speaking-rate, Intensity mean, and Intensity max; similar patterns in convergence at turn level are found in two languages; similar synchrony over turns is found in two languages; over different gender groups, in both languages, the number of prosodic features showing entrainment in the mixed gender pairs is the most, and the least in male pairs.

Differences are also found in entrainment degree in both languages. Convergence degree at turn level is different in two languages. The correlation values, which are used to measure the relation between similarity and time, are different in both languages. Synchrony degree is different in two languages. Correlation values are used to measure synchrony degree

over turns. The most noticeable difference is that in SAE, correlations between two interlocutors' in adjacent IPUs over F0 features are weak, while in MC, they are among the strongest. This indicates that the functions of pitch in MC more prominent in the exhibition of entrainment than in SAE. This proves that pitch in MC plays a dual role in a tonal language, conveying both lexical and pragmatic information. Over different gender groups, in SAE conversations, entrainment is both strongest and prevalent in mixed gender pairs; in MC, it is most prevalent in mixed gender pairs, but not necessarily strongest.

Besides, through the comparison, the absolute entrainment is found in previous research on prosody in English conversations, while both absolute and relative entrainment is proved in the prosody of Mandarin Chinese conversations in the present research, in which absolute entrainment is found over global units, and relative entrainment between the relative pitch registers of tones in the conversations of mixed gender pairs.

The similarity in entrainment pattern in both languages found in the present research supports not only the view that entrainment is a cross-cultural phenomenon but provides evidence that members of different languages groups entrain in similar pattern. Nevertheless, the difference in language types would be one of the reasons for prosodic entrainment degree.

5. MAJOR FINDINGS

To sum up, the major findings of this research are listed as follows:

(i) Entrainment in Mandarin conversation is proved to exist in 3 main aspects of prosody.

The analyses in the present research provide considerable evidences for the existence of entrainment in 3 main aspects of prosody: the features of duration, F0 and intensity. 7 variables, Speaking-rate (one feature of duration), F0min, F0mean, F0max (three features of F0) and Intensity min, Intensity mean, Intensity max (three features of intensity) are set as parameters in the present research, over which entrainment in conversation has been proved.

(ii) For the prosodic features measured in the present research, local entrainment is more evident than global entrainment.

According to the analyses above, more prosodic features exhibit entrainment at relatively local level than at global level. In detail, in terms of proximity, 5 features show entrainment over tone units, and 4 features at conversation level; in terms of synchrony, 7 features show entrainment over tone units, 6 features at the turn level. Besides, the features of F0 show more prominent entrainment at relatively the local level, in spite of their less evident entrainment at the global level.

(iii) Different prosodic features exhibit entrainment at different levels: at both the global and the local levels, features of intensity exhibit prominent entrainment; entrainment over the durational feature is more prominent at the relatively global level; entrainment over the features of F0 is more prominent at the relatively local level.

According to the cross-level comparisons, the features of intensity show entrainment at almost all the levels tested, the features of duration show more prominent entrainment at the conversation level than the turn level, and the features of F0 show evident synchrony at levels of turn and tone unit, while they show less entrainment at the level of conversation.

(iv) With the progress of conversation, increased similarity between two interlocutors is found at the turn level, but no increased similarity is found at the conversation level or over tone units.

The present research has accomplished the analyses of convergence at three levels. The cross-level comparisons of convergence at three levels show that it is more difficult to capture convergence in conversation than proximity and synchrony, and that convergence is found only at turn level.

(v) Entrainment is proved to be closely related to social factors: in terms of gender, prosodic features exhibit differences in entrainment over 3 gender groups, and in both number of features entrained and degree of entrainment, males entrain the least; in terms of role, it is proved that the roles affect the entrainment degree in conversation over the

prosodic features of duration and F0 and that information givers entrain more than followers in conversation.

(vi) Striking similarities in entrainment patterns are found in Mandarin Chinese and Standard American English, and differences in entrainment degree are found in both languages.

The following similarities in entrainment patterns in two languages are found. At levels of conversation and turn, those features, which exhibit holistic entrainment in conversation and those which do not, are the same. At the turn level, those features, which show increasing similarities as the conversation goes on and those which do not, are the same. Similar synchrony over turns is found in two languages. Over different gender groups, in both languages, the number of prosodic features showing entrainment in the mixed gender pairs is the most, and that in male pairs is the least.

The following differences in entrainment degree in both languages are found. At the turn level, differences in convergence and synchrony degree are found in both languages. Over different gender groups, in SAE conversations, entrainment is both strongest and prevalent in mixed gender pairs, while in MC, it is most prevalent in mixed gender pairs, but not necessarily strongest.

(vii) Two kinds of entrainment are proposed in the present research: absolute entrainment and relative entrainment.

Entrainment in previous research on prosody in English conversations is absolute. Both absolute and relative entrainment is proved in the prosody of Mandarin Chinese conversations in the present research, in which absolute entrainment is found over more global units, and relative entrainment is found between mixed gender pairs' relative pitch registers of tones.

6. LIMITATIONS

The present research has limitations, which supply the potential possibilities for the future studies.

Firstly, modeling the entrainment process in conversations should be improved. Capturing dynamic entrainment is difficult. The difficulties exist mainly in two sides. One is that the adaptation of two interlocutors changes with the progress of conversation. These changes might be caused by psychological, social, behavioral, or other factors. It is difficult to find a fixed model to generalize these changes. Another difficulty is that the adaptation partners are not time-aligned. Therefore, it is necessary to find effective methods to model precisely the prosodic entrainment process in conversation in the future studies.

Secondly, more social factors should be involved in the study of prosodic entrainment. The entrainment in conversation involves various social factors of interlocutors: the status, age, gender, role, region, ethnicity, environment features, etc., of which the present research studied the gender and role. In fact, the functions of other social factors are also worth investigating. It is not possible to do research on all the social factors in one study. The difficulty in dealing with these social factors is how to make the target factors prominent and make valid control of other factors in experiments.

Thirdly, the relationship between pitch and prosodic entrainment should be further studied. The present research did a preliminary study, in which prosodic entrainment is analyzed over tone units. Proximity, convergence, and synchrony are tested over all the units. However, for the relationship between tones and prosody in Chinese, majority of literature only supply assumed descriptions rather than explicit explanations. Therefore, it is necessary to make further research on this point.

Author: Xia Zhihua PhD in Tongji University, a lecturer of Jiangsu Normal University. Her research interests include speech prosody, Intonation and L2 language acquisition.
E-mail: xzhlf@163.com

Supervisor: Ma Qiuwu PhD He is a professor of Tongji University. His research interests include phonology, speech prosody, and phonetics.
E-mail: tjnkmqw@126.com

(此博士学位获得 2015 年同济大学优秀博士学位论文称号。)