A Preliminary Study of Prosodic Disambiguation by Chinese EFL Learners

Yuanyuan Zhang, Hongwei Ding*

Institute of Cross-Linguistic Processing and Cognition, School of Foreign Languages,
Shanghai Jiao Tong University, China

yyuan.zhang@sjtu.edu.cn, hwding@sjtu.edu.cn

Abstract

This study investigated whether Chinese learners of English as a foreign language (EFL learners hereafter) could use prosodic cues to resolve syntactically ambiguous sentences in English. 8 sentences with 3 types of syntactic ambiguity were adopted. They were far/near PP attachment, left/right word attachment and wide/narrow scope. In the production experiment, 15 Chinese college students who passed the annual national examination CET (College English Test) Band 4 and 5 native English speakers from America were recruited. They were asked to read the 8 target sentences after hearing the contexts spoken by a Native American speaker, which clarified the intended meaning of the ambiguous sentences. The preliminary results showed that, as the native speakers did, Chinese EFL learners employed different durational patterns to express the alternative meanings of the ambiguous sentences by altering prosodic phrasing. That is, the duration of the pre-boundary items were lengthened and pause were inserted at the boundary. But the perception experiment showed that the utterances produced by Chinese EFL learners couldn't be effectively perceived by the native speakers due to their different use of pre-boundary lengthening and pause. The conclusion is that Chinese EFL learners find prosodic disambiguation difficult.

Index Terms: prosodic disambiguation, syntactic ambiguity, Chinese EFL learners

1. Introduction

Prosody plays an important role in sentence processing, and prosodic phrasing reflects the syntactic structure of the sentence. The presence and absence of prosodic boundaries at the prosodic phrases guide the interpretation of the syntactic structure. Research on the role of prosody in sentence processing and the effect of prosodic phrasing on syntactic structure has primarily focused on the resolution of ambiguity [1]. For example, the sentence "*For our parties, we invite David and Pat or Bob, but not all three*" may indicate that *we invite David, and also Pat or Bob*; or that *we invite both David and Pat, or else Bob* [2]. It's the location of the prosodic boundary that clarifies. Prosodic disambiguation has been thoroughly studied in first language production, perception and processing [2-14]. It's shown that speakers of a language can employ different prosodic cues to resolve ambiguity and convey the alternative meanings of an ambiguous sentence [2-3, 12, 15], and even naïve speakers and listeners reliably use prosodic phrasing to resolve syntactic ambiguity [10, 13]. The prosodic cues characterizing prosodic phrasing include pause, phrase-final lengthening, pitch movement [16], and amplitude

reduction [6]. However, most of the language learners in their early language learning focus on grammatical knowledge, and pay little attention to their prosodic development. Besides, previous studies on the prosodic development of the foreign language learners primarily focus on their rhythm, intonation and stress etc., leaving prosodic phrasing, which is a reflection of the information unit grouping in the utterance, aside. Few studies have been carried out concerning the prosodic disambiguation in second language (L2 hereafter) [17-19].

In the study of the early Spanish-English bilinguals, it's found that the duration of the bilingual is greater than that of the monolingual, and the pitch movement of the bilingual is less extreme than that of the monolingual [20]. English-French L2 learners can use prosodic information for interpretation, but their preference for prosodic phrasing is related to their French proficiency [21]. The advanced Taiwanese-English learners make more distinguishable pause duration difference than the native speakers, but the limited Taiwanese-English learners cannot use pause reliably to disambiguate the sentences [19]. German-English and English-German L2 learners are found not to produce reliable durational differences in prosodic disambiguation, though their pitch movement is similar to that of the native language speakers [22]. It seems that the language learners do have differences from the native speakers, especially in durational pattern, though these studies are carried out in different target languages. It's proposed that the prosodic cues indicating a major prosodic boundary might be language-specific [23]. Thus, it's important to explore whether language learners can use prosodic cues to group the syntactic constituents of a given sentence and express its intended meaning.

Previous studies showed that the prosodic cues for disambiguation in Chinese Mandarin include pause insertion, pre-boundary lengthening, F0 lowering, and intensity reduction, while pause is the primary prosodic cue, and pre-boundary lengthening is less important [23]. While in English, it's the pre-boundary lengthening that primarily marks the prosodic boundary [6]. Due to the possible differences in the use of duration pattern between Mandarin and English, the present study will try to investigate the prosodic disambiguation by Chinese EFL learners, to see the durational pattern that they employ to disambiguate these sentences, and if their prosodic disambiguation can be perceived by the native English speakers (NS).

2. Method

In our study, both production task (Experiment 1) by Chinese EFL learners and perceptual evaluation (Experiment 2) by Native American English speakers were carried out, so that we could know the specific durational pattern that the EFL

---

* Corresponding author

learners employed, and if their prosodic cues could be perceived by the native speakers.

2.1. Experiment 1: production task

2.1.1. Recording materials

In this study, the recording materials included 8 sentences with 3 types of syntactic ambiguity adopted from previous studies [24]. They were far/near prepositional phrase (PP hereafter) attachment, left/right word attachment and wide/narrow scope. All the target sentences were inserted in a corresponding context, which illustrated the intended meaning of the ambiguous sentences. Those interpreted with a prosodic boundary between the critical constituents were grouped into context A, while those without a boundary into context B. The context length varied from one to four sentences, and the target sentences were inserted as the final part. The contexts and the target sentences (italic) are displayed in Table 1.

Table 1: *Recording materials*

1. Far/near attachment
(1A) The girl's favorite English teacher left. When she saw the new teacher, she seemed very unhappy. *The teacher greeted the girl with a smile.*
(1B) The teacher didn't want to greet anyone because he felt unhappy yesterday. When the teacher came into the classroom, he found a girl was smiling at him. *The teacher greeted the girl with a smile.*
(2A) Nick's defense lawyer claimed that the murder weapon was a knife, but when we saw the body we knew they were wrong. *Nick murdered the man with a gun.*
(2B) Nick denied murdering either the man who carried a gun or his unarmed accomplice. We don't know yet who murdered the unarmed accomplice, but Nick's bloodstained knife clinched the case for the other murder. *Nick murdered the man with a gun.*
(3A) I had no time to change clothes when my friend came to visit me on Sunday morning. *I saw my friend in my pajamas.*
(3B) One of my friends visited me and stayed here for a week. When I came back at night during her visit, she opened the door for me and I was astonished. *I saw my friend in my pajamas.*
2. Left/right word attachment
(4A) Bears sleep all winter long, usually coming out of hibernation in late April, but this year they were a little slow. *They rose early in May.*
(4B) In spring there was always more work to do on the farm. May was the hardest month. *They rose early in May.*
(5A) Every time Lynda comes to the canteen to have lunch, she chooses a seat by the window side. But yesterday another man occupied that seat. There, *she saw a man eating fish.*
(5B) The ship was sailing before the wind. She took a telescope to enjoy the broad view of the sea. Suddenly she cried out because a huge fish jumped out of the water. There, *she saw a man eating fish.*
(6A) Anna likes meat very much. Unfortunately, she lost her sense of taste in a traffic accident last week. Now, Anna

doesn't know how good meat tastes.
(6B) Anna's mother works in the restaurant. Every time Anna wants to have meat, her mother will bring some leftover meat to her from the restaurant. *Anna doesn't know how good meat tastes.*
3. Wide/narrow scope
(7A) Taylor' dogs are so small that they seem to be just born, but the cats seem to be big enough to give birth to baby cats. So, *Taylor must raise tiny dogs and cats.*
(7B) One of Taylor's friends' dogs and cats gave birth to baby dogs and cats last week. Taylor likes them so much that he wants to take them home. So, *Taylor must raise tiny dogs and cats.*
(8A) When the war broke out, the young man went to the battlefield. Those who were not strong enough were allowed to stay at home. So, *the old men and women stayed at home.*
(8B) In the past decade, the young people in the countryside went to work in the city. So, *the old men and women stayed at home.*

Each of the short paragraph was followed by a contextualized question, for example, the question followed (1A) was "*Who was smiling?*" to draw the participants' attention to the intended meaning of the sentence. Before the experiment, all the short paragraphs and the contextualized questions were read aloud and recorded in a random order by a Native American English speaker (NS hereafter) in a natural and fluent way for the pilot study and the production task.

In the pilot study, 4 Chinese EFL learners majored in English, a phonetician and a native British English speaker were presented with the target sentences and the questions read by the NS, and meanwhile presented with the two possible contexts for the target sentences in written form. They were asked to choose a context that best suited the target sentence according the question and the prosody of the sentence that they heard. It turned out that all of them could choose the appropriate contexts for the target sentences. But they needed listen to the recordings repeatedly to make decisions. It indicated that the NS disambiguated all the target sentences prosodically, and we could employ the same experiment paradigm in the production task.

After the pilot study, the recording materials for the production task were divided into two parts, each part with half context A and half context B, but no sentences with two alternative meanings in one part.

2.1.2. Subjects

The subjects were grouped into EFL learners and the native speakers (NS). 15 Chinese EFL learners (8 male, 7 female) from Shandong, China, and 5 Native American English speakers (4 male, 1 female) were recruited. All of the EFL learners were college students from Shanghai. They spoke Shandong dialect and Mandarin. They have passed the annual national examination CET (College English Test) Band 4. The NS were also college students who came to China for short-term international exchange. They aged from 19 to 24. None of them reported any speech or hearing disorders.

2.1.3. Acoustic measurement

The recording took place in the professional recording studio at the School of Foreign Languages in Shanghai Jiao Tong University. The speech was digitized into a computer (Macbook Pro) by 16Bit/44.1K using audio sound (AVID Mbox Mini) and microphone (Sennheiser). The subjects sat comfortably in front of the computer, with the microphone being placed approximately 5 cm away from their lips. The recording materials were presented to the subjects in auditory and written form. Before recording, the subjects would have enough time to be familiar with the recording materials, to ensure they knew each word and had a good understanding of the sentences. For each sentence recording, the subjects were asked to read the context and the target sentence silently on the computer screen, and then they would hear the context read by the NS. They were asked to hear the contexts carefully to get its intended meaning, and after that to read aloud the target sentence in a natural and fluent way to convey its intended meaning using prosodic information. Each EFL speaker recorded each target sentence twice and NS once in separate randomized blocks generated by Speechrecorder [25]. If any mistake was judged by the experimenter during the recording, such as mistakes in segmental pronunciation, or adding or deleting any word, the participants would be asked to repeat the sentences, which would only happened occasionally. Besides, the experimenter would ask the contextualized question randomly to ensure that the participants understood the context and the target sentence. Totally, 480 utterances were produced by EFL speakers, and 80 utterances by NS. For EFL subjects, only sentences recorded for the second time were used. Thus, 320 utterances would be analyzed in this study.

2.1.4. Acoustic measurement

In the present study, the acoustic measurement was done with Praat [26]. It focused on the critical items that straddled the site for an additional syntactic boundary [24]. The critical items were underlined in table 1. The data extraction was done with a Praat script ProsodyPro [27]. Every syllable of the utterances was segmented and labeled as an interval. The irregular or missed pulse marking were corrected. After running the script, the duration of the critical item and the pause between the pre- and post-boundary items were extracted and measured.

Individual differences in speaking rate were normalized across speakers by taking the duration ratios of the pre-boundary critical items and pause. The duration ratio was derived by dividing the duration of the critical items or pause from that of the whole sentence. Thus, the pre-boundary duration ratio (PreDR) and pause duration ratio (PDR) were finally obtained.

2.1.5. Results

Figure 1 displays the mean PreDR and the mean PDR in context A (PreDRA, PDRA) and context B (PreDR, PDRB). 1-8 on the Y-axis means the 8 sentences read by NS, and 9-16 refers to those read by EFL learners.

From Figure 1, we can see that when the target sentences are interpreted in context A, in which a prosodic boundary should be inserted, the PreDR is greater than the same constituents in context B for NS, except sentence 3 in which the PreDR in context A is smaller than that in context B. This

pre-boundary lengthening effect by NS is similar to that by EFL learners. But the duration ratio difference between the same items of the sentence in the two contexts by EFL learners is smaller than that by NS. For the pause effect, it's found that except in sentence 4, there is no pause in context B for NS, but pause exists in both contexts for EFL learners.
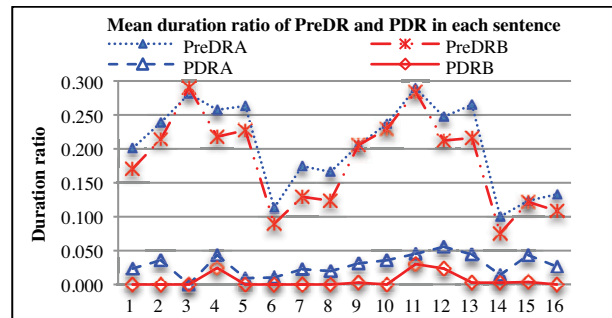

Figure 1: *Mean duration ration of the pre-boundary critical items and pause in each utterance by NS (1-8) and EFL (9-16).*

To determine if the prosodic boundary effect on pre-boundary lengthening and pause is significant for NS and EFL learners, and if the difference between the two groups is significant, a two-way ANOVA was carried out, with the factor context (context A and context B) as the within-subjects variable and group (NS and EFL learners) as the between-subjects factor. Table 2 shows the result of the ANOVA analysis.

Table 2: *The F and p values of average PreDR and PDR averaged across 8 sentences by Mixed Between-Within ANOVA to determine the effect of context and group.*

Dependent variables	Effect	df	F	p
PreDR	Context	(1,18)	8.429	**0.009**
	Group	(1,18)	1.381	0.255
	Context*Group	(1,18)	1.832	0.193
PDR	Context	(1,18)	16.702	**0.001**
	Group	(1,18)	3.074	0.097
	Context*Group	(1,18)	0.716	0.409

It can be seen from the above ANOVA analysis that context have a significant effect on the PreDR and PDR for both groups. The duration of the pre-boundary items is lengthened and pause is inserted between pre- and post-boundary items in context A meaning. But the interaction between context and group is not significant, indicating that the difference in the context effect on PreDR and PDR is not significant between the two groups.

2.2. Experiment 2: perceptual evaluation

2.2.1. Perception materials

The target utterances produced by the EFL learners for the second time in the production task were collected. They were grouped into 4 blocks randomly, and then were programmed into E-prime (2.0) with contexts. The contexts were presented to the subjects in written form, and the target sentences were audio sounds. For each audio sound, there were three choices for the subjects: "context A", "context B" and "not sure". The

audio sounds were randomly presented to the subjects, and the answers for "A" and "B" were balanced.

2.2.2. Subjects

After finishing the production task, 4 (3 male, 1 female) of the NS subjects in experiment 1 took part in the perception task after a short rest.

2.2.3. Procedure

The subjects took a comfortable seat in front of the computer. They were told to judge, using prosodic information, which context corresponded to the utterance that they heard by pressing "1", "2", "3" on the key. The audio sound would be played repeatedly until they pressed one of the three keys. They would have enough time to respond in this way.

2.2.4. Result

Table 3: *The accuracy rate of the utterances produced by EFL learners judged by NS.*

Speaker	A	B	Total
Sp1	**0.375**	0.875	0.625
Sp2	**0.375**	**0.375**	0.375
Sp3	0.625	0.750	0.688
Sp4	0.625	0.625	0.625
Sp5	**0.500**	0.750	0.625
Sp6	0.625	0.625	0.625
Sp7	**0.375**	0.625	0.500
Sp8	0.625	**0.500**	0.563
Sp9	0.625	0.750	0.688
Sp10	0.625	**0.500**	0.563
Sp11	**0.500**	**0.500**	0.500
Sp12	**0.500**	**0.250**	0.375
Sp13	0.750	**0.500**	0.625
Sp14	0.625	0.625	0.625
Sp15	**0.375**	**0.500**	0.438

Table 3 presents us the accuracy rate judged by NS listening to sentences produced by EFL learners. The column "Speaker, A, B and Total" in Table 3 represents the serial number of the EFL subjects who produced the sentences in experiment 1; the intended meaning in context A; the intended meaning in context B; and the overall accuracy rate of the sentences in both context A and context B.

The total accuracy rate varied from 37.5% to 68.8%. For context A meaning, it ranged from 37.5% to 75%; for context B meaning, 25% to 87.5%. We picked out the EFL learners' production data of those accuracy rate≤ 50% and carried out the Independent-Samples T test to see if the difference in the durational pattern between NS and EFL learners had significance. The result showed that for context A meaning, the mean PreDR of NS is significantly higher than that of EFL learners (0.200>0.029) (p=0.013, <0.05), but no significant effect was found on the PDR (p=0.305, >0.05) between the two groups, though the mean PDR of EFL learners is higher than that of NS (.029>.017). In the interpretation of context B meaning, the T test showed no significance on PreDR (p=0.076, >0.05) or PDR (p=0.271, >0.05), though the mean PreDR (.186) and the mean PDR (.005) of NS is lower than that of the EFL learners (.188, .013).

It's showed that duration plays a greater role in interpreting the structurally ambiguous sentences when duration and pitch contour are pitted against one another [6, 28] found that. Thus, from the above analysis, we can infer that it

might be the difference in durational pattern between NS and EFL learners that lead to the lower accuracy rate.

3. Discussion and conclusion

This is a primary investigation into Chinese EFL learners' production of prosodic disambiguation. It's found that the alternative meanings of the ambiguous sentences exerted significant effects on PreDR and PDR for both NS and EFL learners. Specifically, when the sentences were interpreted with meaning A in which a boundary was inserted between the critical items, pre-boundary duration was lengthened and pause was inserted. But statistic analysis showed no significant differences between NS and EFL learners. In the perception task, accuracy rate of sentences produced by some of the EFL speakers' was found to be ≤ 50%. The T test showed that the mean PreDR of the sentences interpreted with meaning A by NS was significantly longer than that by EFL learners whose accuracy rate≤ 50%. Thus, it's inferred that Chinese EFL learners' use of pause was more distinguishable than pre-boundary lengthening to mark prosodic boundary, which might cause misinterpretation by NS. As for the insignificant difference in the mean PDR between the NS and EFL learners, it's suggested that the NS made judgments according to the overall prosodic information of the whole utterance, so pitch movement and prosodic cues of other constituents in the utterances might influence the accuracy rate.

The study of Taiwanese-English [19] found that the advanced Taiwanese-English learners make more distinguishable pause duration difference than the native speakers, and the limited English learners don't create a significant difference tendency in the pause duration. While Chinese EFL learners use significantly shorter pre-boundary lengthening than NS. The distinction in the use of durational patterns might be the reason why the learners' prosodic disambiguation couldn't be effectively perceived by NS, though they altered prosodic phrasing when expressed the alternative meaning of the ambiguous sentences.

In conclusion, this study presents us that both NS and Chinese EFL learners employed different durational patterns to express the alternative meaning of the ambiguous sentences by altering the prosodic phrasing. But the perception task indicated that the learners' prosodic disambiguation was not effective. It's assumed that some of the learners relied more on pause and might overuse pause at places where no prosodic phrasing was needed.

Some limitations still exist in our study. First, only the critical constituents of the pre- and post-boundary were analysed. Second, only durational pattern was analysed in this study, leaving pitch change aside, which has been proved to be an important prosodic cue to mark prosodic boundary. Finally, we didn't group the EFL learners into different proficiency levels which have been proved to be a factor in the prosodic disambiguation. The above factors mentioned in the limitations will be explored in further studies.

5. References

[1] H. N. Nagel, L. P. Shapiro, B. Tuller, and R. Nawy, "Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing," *J Psycholinguist Res,* vol. 25, pp. 319-44, 1996.

[2] D. W. Allbritton, G. McKoon, and R. Ratcliff, "Reliability of prosodic cues for resolving syntactic ambiguity," *J Exp Psychol Learn Mem Cogn,* vol. 22, pp. 714-35, 1996.

[3] I. Lehiste, "Phonetic disambiguation of syntactic ambiguity," *J Acoust Soc Am,* vol. 53, p. 380, 1973.

[4] D. R. Scott, "Duration as a cue to the perception of a phrase boundary," *J Acoust Soc Am,* vol. 71, pp. 996-1007, 1982.

[5] P. J. Price, M. Ostendorf, S. Shattuck-Hufnagel, and C. Fong, "The use of prosody in syntactic disambiguation," *J Acoust Soc Am,* vol. 90, pp. 2956-70, 1991.

[6] L. A. Streeter, "Acoustic determinants of phrase boundary perception," *J Acoust Soc Am,* vol. 64, pp. 1582-92, 1978.

[7] J. Pynte, "Prosodic breaks and attachment decisions in sentence parsing," *Language and Cognitive Processes,* vol. 11, pp. 165-192, 1996.

[8] A. J. Schafer, "Prosodic parsing: The role of prosody in sentence comprehension" [D]. University of Massachusetts Amherst, 1997.

[9] A. J. Schafer, S. R. Speer, P. Warren, and S. D. White, "Intonational disambiguation in sentence production and comprehension," *J Psycholinguist Res,* vol. 29, pp. 169-82, 2000.

[10] A. J. Schafer, S. R. Speer, and P. Warren, "Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task," *Approaches to studying world-situated language use,* pp. 209-225, 2005.

[11] M. M. Kjelgaard, and S. R. Speer, "Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity," *Journal of Memory and Language,* vol. 40, pp. 153-194, 1999.

[12] J. Snedeker, and J. Trueswell, "Using prosody to avoid ambiguity Effects of speaker awareness and referential context," *Journal of Memory and language,* vol. 48, pp. 103–130, 2003.

[13] S. R. Speer, P. Warren, and A. J. Schafer, "Situationally independent prosodic phrasing," *Laboratory Phonology,* vol. 2, pp. 35-98, 2011.

[14] K. Steinhauer, K. Alter, and A. D. Friederici, "Brain potentials indicate immediate use of prosodic cues in natural speech processing," *Nat Neurosci,* vol. 2, pp. 191-6, 1999.

[15] S. R. Speer, M. M. Kjelgaard, and K. M. Dobroth, "The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities," *Journal of psycholinguistic research,* vol. 25, pp. 249-271, 1996.

[16] J. R. D. Pijper, and A. A. Sanderman, "On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues," *J Acoust Soc Am,* vol. 96, pp. 2037-2047, 1994.

[17] H. Hwang and A. J. Schafer, "Prosodic effects in parsing early vs. late closure sentences by second language learners and native speakers," *Proceedings of the 3rd International Conference on Speech Prosody,* pp. 585-88, 2006.

[18] L. Dekydtspotter, B. Donaldson, and A. C. Edmonds et al., "Syntactic and prosodic computations in the resolution of relative clause attachment ambiguity by English-French learners," *Studies in Second Language Acquisition,* vol. 30, pp. 453-480, 2008.

[19] P. L. Yang, "English Language Proficiency and Production of Prosodic Disambiguation A Preliminary Study of Taiwanese English Learners," *Journal of English Teaching (英語教學期刊),* vol. 34, pp. 51-84, 2010.

[20] E. M. Fernández, "The prosody produced by Spanish-English bilinguals A preliminary investigation and implications for sentence processing," *Revista da ABRALIN,* vol. 4, pp. 109-141, 2005.

[21] A. L. Fultz, "The Use of Prosody for Disambiguation in English-French Interlanguage," *9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007). Cascadilla Proceedings Project,* pp. 130-139, 2008.

[22] M. G. O'Brien, C. N. Jackson, and C. E. Gardner, "Cross-linguistic differences in prosodic cues to syntactic disambiguation in German and English," *Appl Psycholinguist,* vol. 35, pp. 27-70, 2014.

[23] X. S. Shen, "The use of prosody in disambiguation in Mandarin," *Phonetica,* vol. 50, pp. 261-271, 1993.

[24] W. K. Choe and M. A. Redford, "Assumed Meaning in the Production and Use of Disambiguating Prosody," spl.uoregon.edu

[25] "SpeechRecorder", was originally presented at the International Conference on Language Resources and Evaluation in Lisbon in 2004 [DJ04], 2004.

[26] P. Boersma and D. Weenink, Praat: doing phonetics by computer [computer program]. Version 6.0.23. retrieved 24 December, 2016. [Online]. Available: http://www.praat.org

[27] Y. Xu. (2005-2017) Prosodypro.praat. [Online]. Available: http://www.phon.ucl.ac.uk/home/yi/ProsodyPro/

[28] C. M. Beach, W. F. Katz, and A. Skowronski, "Children's processing of prosodic cues for phrasal interpretation". *J Acoust Soc Am*, vol. 99, pp. 1148– 1160, 1996.