Prosodic Grouping and Relative Clause Disambiguation in Mandarin

Jianjing Kuang

Department of Linguistics, University of California Los Angeles, U.S.A.

`kuangjianjing@ucla.edu`

Abstract

The study discusses the role of prosodic grouping in the Mandarin Relative Clause attachment disambiguation. The grouping effect is explored under the Implicit Prosody Hypothesis (IPH) from four aspects of sentence processing experiments: default production, contrast production, online processing, as well as offline processing. It is found that (1) the length of RC greatly impacts ambiguity resolutions offline; (2) Prosodic grouping can well reflect the different attachment readings and is consciously used to produce contrastive meanings (3) Online processing can be affected by manipulating the grouping cues: Prominence and pause. The findings support the IPH, and contribute to our understanding about prosodic grouping in Mandarin, which can be applied in spoken language processing.

Index Terms: prosodic grouping, disambiguation, Mandarin

1. Introduction

Syntactic ambiguity arises when a given utterance can be represented by more than one syntactic structure, bringing the difficult problem to spoken language processing. One of the frequently studied cases is Relative Clause (RC) attachment ambiguity. The classical example is "*the servant of the actress who was on the balcony*". In this sentence, a complex NP head is modified by a relative clause and the relative clause can be ambiguously attached either to the "high" NP ("*servant*") or to the "low" NP ("*actress*"). Given no semantic and syntactic bias in the attachment, it has been found that the preference of RC attachment differs across languages [1][2].

The Implicit Prosody Hypothesis (IPH) ([3],[4]) states that the preference difference attributes to the different prosody across languages. In silent reading, default prosodic parsing is projected onto the sentence and influences syntactic ambiguity resolution. Overt prosody is assumed to be the same as implicit prosody. Moreover, the IPH predicts that a prosodic grouping of (NP1) (NP2 RC) indicates a low attachment interpretation (i.e., RC attaches to NP2), while a prosodic grouping of (NP1 NP2)(RC) reflects a high attachment interpretation (i.e., RC attaches to NP1). Studies ([1][6]) have shown that prosodic grouping can affect processing both offline and online.

The goal of the present study is to extensively examine how Mandarin speakers prosodify this type of complex NP phrases and whether the prosodic grouping is used for disambiguation of RC attachment in Mandarin. This is done by testing the proposals stated by IPH. Unlike many other investigated languages, Mandarin is a tonal language and thus has distinctive prosodic properties. In addition, Mandarin like Korean and Japanese has a RC NP1 NP2 word order, which is the opposite order of English and less common among the investigated languages [2]. So data from Mandarin is very important for the validation of IPH. Other than the theoretical goal, the experimental facts are meaningful too. The knowledge in terms of when prosodic grouping matters and how it realized can help to solve the difficult problems in spoken language processing.

According to the IPH, grouping effect can be explored based on the following hypotheses, which reflect the different aspects of sentence processing:

- Manipulating the prosodic factors (i.e. phonological length in this study) could affect offline ambiguity resolution; (implicit prosody)
- Default produced prosody matches with the offline sentence interpretation.(default production)
- Speakers are able to consciously make contrastive prosodic groupings to remove ambiguity (contrast)
- Listeners are able to perceive the different groupings by manipulating the prosodic features. (auditory perception)

Four experiments are conducted with respect to each hypothesis. Grouping preference in silent reading, in default production, in contrast production and auditory perception of grouping are examined.

2. Experimental materials

The frame of target sentences is "RC de NP1 de NP2". Examples:

(1)Zhanzai jiangtai shang DE   jiaoshou DE xuesheng
   Who stands on the stage DE professor DE student
   'The student of the professor who stands on the stage'

(2)Huojiang   DE   jiaoshou DE   xuesheng
   Who got the prize DE professor DE student
   'The student of the professor who got the prize'

Eight types of target sentences are created by manipulating the length of each constituent, RC, NP1 and NP2. The length varied in two categories: short and long. Short versions contain 2 syllables (one standard foot in Mandarin), while long versions have 5 syllables (longer than two feet). A full combination of Length types in each constituent is shown in Table 1.

Table1. *Eight types of sentences in varying the length (S: short and L: long) of each constituent, RC, NP1 and NP2*

TYPE	RC	NP1	NP2
1	L	S	S
2	L	S	L
3	L	L	S
4	L	L	L
5	S	S	S
6	S	S	L
7	S	L	S
8	S	L	L

In addition to 8 target sentences, 16 fillers are added. Fillers share the same structure and type with the target sentences, but unlike the target sentences, all the fillers have unambiguous interpretations according to pragmatic or semantic context of the sentences, specifically, eight with low attachment meaning

and eight with high attachment meaning. They are made as the control condition to balance the bias in processing.

3. Preference survey

3.1. Procedures

This preference survey is to investigate if there is attachment preference of RC by native speakers of Mandarin. Thirty native Mandarin speakers participated in the preference survey. A questionnaire is made by pseudo-randomizing the 8 target sentences and 16 fillers sentences. Informants were asked to silently read the sentences, which were written in Chinese characters, and asked to answer one of the two questions "Who was on the stage/ who got the prize?" They circled either 'professor' or 'student' on the sheet. This was also done through emails.

3.2. Results

The results show that native Mandarin speakers prefer low attachment regardless of the prosodic (length) conditions. As shown in Figure1, about 70% of the choices were NP1, in a sequence of RC NP1 NP2.
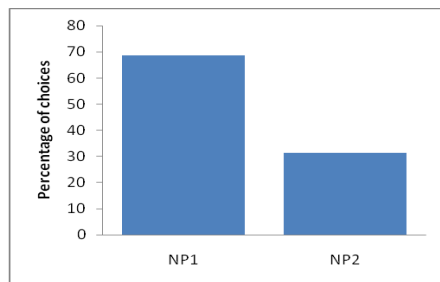

Figure 1: *Attachment preference in the RC NP1 NP2 structure.*

The choices were then sorted by the length of RC. As shown in Figure 2, a clearly unbalanced pattern is found. Fisher exact test shows that RC length has significant influence on the choices (p<0.001).
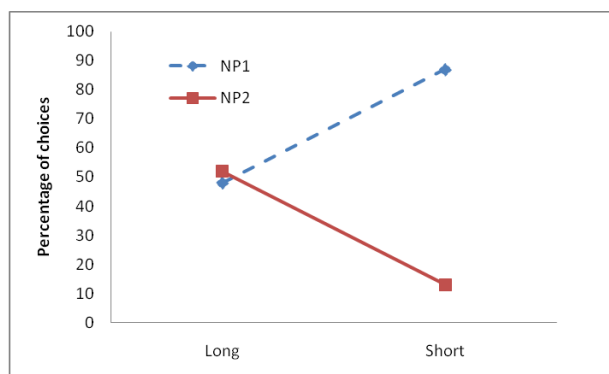

Figure 2: *Attachment preference (NP1 vs. NP2) in each Length condition of RC*

When RC is long, choices between NP1 and NP2 are quite similar; but when RC is short, choices of NP1 overwhelms the choices of NP2. In other words, short RC leads to strong preference of low attachment. The same strategy was applied for the case of NP1 and NP2 length, but no significant effect

was discovered. This result partially supports about the effect of phonological length in attachment [1][3][4][6][7].

4. Productions

Two types of production experiments were conducted to investigate the default prosody of the target structure when reading out-of-the-blue and the prosody employed when NP1 or NP2 attachment meaning is forced, i.e., contrast reading.

4.1. Procedures

4.1.1. Default reading

The purpose of the experiment is to examine the most natural (default) prosody of the target structure for the Mandarin speakers. The speakers were asked to produce the 16 target sentences as well as 16 fillers according to their understanding of the sentences. All the sentences are labeled in *Praat*, following the M_ToBI convention [5].

4.1.2. Contrast reading

The goal of this experiment is to explore the prosodic parsing strategy of the Mandarin speakers. When the speakers are forced to make a contrast between high attachment reading (RC) (NP1 NP2) and low attachment reading (RC NP1) (NP2), they have to employ some salient prosodic cues to indicate the different groupings. The strategies used by native speakers can give us better understanding of prosodic boundaries of Mandarin. In this experiment, for each target sentence, speakers were asked to produce the contrastive meanings in the way of clarifying misunderstanding. The contrast pairs were carried in the "clarification" frame "*wo de yisi shi … er bushi* (I mean… but not…)". All the sentences are labeled following the M_ToBI convention [5].

4.2. Results

The prosodic grouping is defined by the relation among possible breaks. In the sequence of RC NP1 NP2, assuming the prosodic break between RC and NP1 is 'pi' and that between NP1 and NP2 is 'p2', i.e., RC $_{p1}$ NP1$_{p2}$ NP2, an early boundary is when *p1>p2* while a late boundary is when *p1<p2*. And *p1=p2* means no sub-grouping among the three components (i.e., neutral boundary). We expect an early boundary for high attachment reading and a late boundary for low attachment reading.

Figure3 shows that the prosodic phrasing generated in out-of-the-blue default reading and sentence interpretation in default reading do not perfectly match. About half of the sentences are produced without sub-grouping (i.e. 'neutral' phrasing), but the sentences produced with sub-grouping do reveal good correspondence between the prosodic structure and the attachment resolution. More late boundary phrasing was interpreted as low attachment while more early boundary was interpreted as high attachment. This shows that sub-grouping is not obligatory in default reading, but it is correlated with sentence interpretation.

Figure 4 shows the results of phrasing when the subjects were put into the contrast context to indicate different meanings. As expected, in contrast reading, consistent patterns for the distinctive prosodic grouping are displayed for the two attachment readings. Two common strategies are found from the current data.
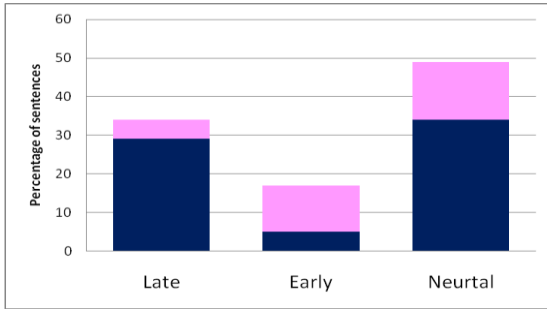
Figure 3: *General preference.*

(1) Prominence contrast [1]

Prominence relation among the three components can lead to different groupings. If it is a RC//NP1 NP2 grouping, the RC is relatively weak and NP1 is the first prominent unit of the structure and more prominent than NP2; if the grouping is RC NP1// NP2, then the RC is accented, more prominent than NP1. So this shows that prominence can mark the beginning of the juncture. The grouping due to contrastive meaning is easier to be seen in the intensity contours. In Figure 4, the first arrow marks the beginning of the RC and the second arrow marks the beginning of NP1. The solid arrow indicates the prominent position.

Figure 4. Prominence relation patterns:
(Zhan51zai51jiang21tai35shang51de)(yu21yan35xue35jiao51 shou51de)(bo35shi51yan35jiu55sheng55)
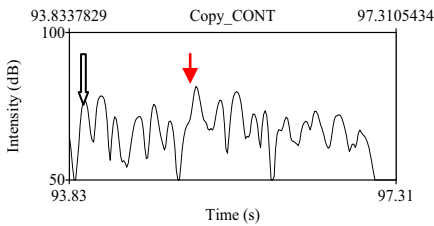"The Ph.D student of the linguistic professor who stands on the stage"

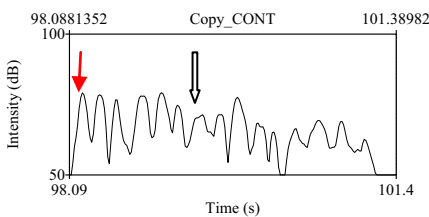Figure 4a: *High attachment reading—NP1 is more prominent than RC.*

Figure 4b: *Low attachment reading— RC is more prominent than NP1.*

(2) Pause

---

[1] So far we avoid using "pitch accent" and "focus", but using the more general term "prominence", since we still need more evidence to convince the domain of the prominence. "Prominence" here is defined as loudness, usually cued by enlarged pitch range and strong intensity.

Not surprisingly, pause is another common strategy. If the grouping is RC//NP1 NP2, then a big early boundary can be found; if the grouping is RC NP1// NP2, a big late boundary can be observed.

Therefore, a strong correlation between prosodic grouping and attachment resolution is found in the contrast reading. In this context, speakers seem to exaggerate and magnify the prosodic cues which are weak or hidden in the default reading.

5. Perception

5.1. Procedures

The purpose of the experiment is to study how prosodic cues contribute to the perception of parsing. From the production experiment, we learned that the location of pause and prominent word is effective prosodic cues in indicating groupings. If the two prosodic cues are manipulated, will the perception of grouping change accordingly? For this purpose, the location of pause and prominence was paired in four ways: PAUSE (early, late) × PROM (NP1, NP2). Thus four versions of stimuli are made for each target sentence. There are 32 stimuli in total.

Examples of stimuli (bolded and capital letters indicate prominence; "//" indicates pause):

(1) RC// **NP1** np2:
Zhanzai jiangtai shang DE //JIAOSHOU DE xuesheng
Who stands on the stage   DE// **professor** DE   student
"The student of the professor who stands on the stage"

(2) RC//np1 **NP2**:
Zhanzai jiangtai shang DE// jiaoshou DE XUESHENG
Who stands on the stage   //DE professor DE   **student**
"The student of the professor who stands on the stage"

(3) RC **NP1**//np2
Zhanzai jiangtai shang DE JIAOSHOU DE// xuesheng
Who stands on the stage   DE **professor** DE // student
"The student of the professor who stands on the stage"

(4) RC np1// **NP2**
Zhanzai jiangtai shang DE jiaoshou DE //XUESHENG
Who stands on the stage   DE professor DE // **student**
"The student of the professor who stands on the stage"

The stimuli are natural signals, produced by a trained phonetician.

Speakers are asked to listen to the stimuli and answer the diagnostic questions, "Who was on the stage/who got the prize". They would circle either "professor" or "student" on the answer sheet. All the tokens are heard for five times.

5.2. Results

Figure 5 shows the percentage of NP1 and NP2 preference when the sentences vary by the location of pause and the prominent word across eight speakers.
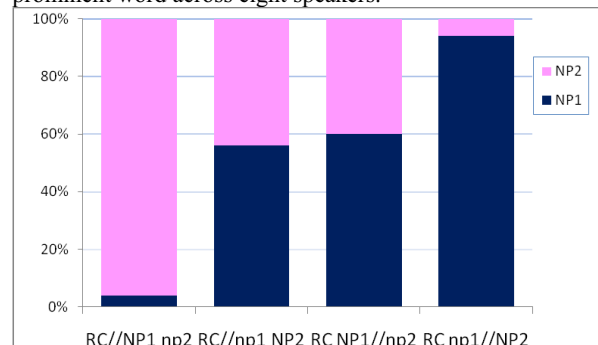
Figure 5: *Percentage of NP1 or NP2 preference given four combinations of PAUSE and PROM, across four speakers.*

Prominent units are marked by capital letters and pauses are marked by "//". Some trends have been found, though the current sample size is small.

As seen in Figure 5, in RC//**NP1** np2 condition, most of the time speakers preferred high attachment while they preferred low attachment in RC np1//**NP2** condition. This trend suggests that (1) pause at early boundary leads to high attachment while pause at late boundary leads to low attachment; (2) Prominence marks the beginning of the grouping so when NP1 is more prominent than NP2, high attachment is preferred but when NP2 is more prominent than NP1, low attachment is preferred. When the two cues cooperate with each other, high agreement of preference is found; but if the two cues conflict with each other (i.e. RC//np1 **NP2**, RC **NP1**//np2), subjects get confused. Subjects also showed individual preferences. When these cues conflict, some subjects put more weight on the pause cue while the others on the prominence cue.
.

6. General discussion

This study examined the prosody of the RC NP1 NP2 structure in Mandarin and the role of prosodic grouping in RC attachment disambiguation. It is found that Mandarin speakers generally prefer a low attachment, and that the attachment resolution is strongly influenced by the length of RC (i.e., longer RC sentences lead to high attachment.)

The IPH claims that different RC attachment preferences across languages are due to the different prosody of languages. But it is unclear what prosodic features should be considered in the comparison. Based on our finding, different RC length across languages might be one of these prosodic factors. The supporters of the IPH have also predicted that the attachment matches the prosodic phrasing, so it is expected that Mandarin speakers would produce (RC NP1)(NP2) more often. However, a good match between the phrasing and attachment was not found in default reading. Only the sentences produced with sub-grouping showed good correspondence to attachment resolution. It suggests that speakers might be conscious of prosodic grouping, but in default reading, they produce the groupings optionally. Then when the context requires a contrast between NP1 and NP1, the optional prosodic features turn to be salient and magnified, and show stronger correlation between prosodic grouping and sentence interpretation.

This study also shows that Mandarin has at least two ways to cue grouping. In addition to salient pause at the grouping boundary, prominence relation among the components also plays a crucial role in grouping. RC is the most prominent in the low attachment grouping (RC NP1//NP2) while it is less prominent than NP1 in the high attachment grouping (RC//NP1 NP2). Essentially, prominence can mark the left edge of a prosodic juncture.

Finally, the perception experiment shows that the pause and prominence both can effectively indicate grouping. More importantly, they collaborate with each other to enhance the auditory cues to a prosodic boundary. When these two cues are compatible with each other, RC attachment preference reaches to the highest agreement level across speakers. By contrast, subjects tend to be confused about grouping when those cues conflict. This finding provides insight into the realization of grouping. Most studies on prosodic grouping have focused primarily on prosodic breaks [8][9]. But prominence has been rarely examined as a boundary cue. Prominence is especially important for Mandarin because the language has no tonal mark for a small prosodic phrase such as iP and AP and has not been analyzed in terms of post-lexical pitch accents. In other words, the Intonation tier of M_ToBI can contribute very little to prosodic grouping. Our finding reveals that prominence is an additional important prosodic grouping cue other than breaks for Mandarin, which is comparable with the previous studies in other languages[10][11]. It suggests that prominence relation is another effective prosodic grouping cue. This knowledge can absolutely benefit spoken language processing models.

7. Conclusion

The study discussed the role of prosodic grouping in the Mandarin RC attachment disambiguation in both production and perception. The IPH is indirectly supported by the facts that (1) the length of RC greatly impacts offline ambiguity resolution; (2) Prosodic grouping can well reflect the different attachment readings; and (3) it is consciously used to deliver contrastive meanings. The perception data show that online processing can be affected by the manipulation of the grouping cues: Prominence and pause. The findings in the study not only contribute to the prosodic universal, but also make us better understand prosodic grouping in Mandarin.

8. References

[1] Jun, S.-A., "Prosodic Phrasing and Attachment Preferences", J. of Psycholinguistic Research, 32(2):219-49, 2003.

[2] Miyamoto, E.. Table listing experimental results of relative clause attachment preferences, 2001.available at http://etm4rc.googlepages.com/table.html.

[3] Fodor, J. D., "Learning to Parse", J. of Psycholinguistic Research, 27(2): 285-319, 1998.

[4] Fodor, J. D., "Prosodic Disambiguation in Silent Reading", NELS , 32: 113-132, 2002.

[5] Tseng C., and Chou F. "A Prosodic Labeling System for Mandarin Speech Database", Proceedings of ICPhS'99,San Francisco, USA, 1999

[6] Wijnen, F. "The implicit prosody of Jabberwocky and the relative clause attachment riddle." In H. Quene & V.J.van Heuven (eds), On speech and language: Studies for Sieb G.Nooteboom. Utrecht: Netherlands Graduate School of Linguistics, pp. 169 -178. 2004

[7] Fernández, E. M. and D. Bradley., "Length effects in the attachment of relative clauses in English", Poster presented at the 12th Annual CUNY Conference on Human Sentence Processing, New York, 1999.

[8] Cutler, A., Dahan, D., & van Donselaar, W. Prosody in the comprehension of spoken language: A literature review. Language and Speech, 40, 141-201.1997

[9] Carlson, K., Clifton, C. and Frazier, L., "Prosodic boundaries in adjunct attachment", Journal of Memory and Language, 45(1): 58-81, 2001

[10] Quinn, D., Abdelghany, H. and Fodor, J.D. , "More evidence of implicit prosody in reading: French and Arabic relative clauses", Poster presented at the 13th Annual CUNY Conference on Human Sentence Processing, La Jolla, CA, March 30-April 1, 2000.

[11] Maynell, L. A., "Effect of pitch accent placement on resolving relative clause ambiguity in English", Poster presented at the 12th Annual CUNY Conference on Human Sentence Processing, New York, 1999.