THE ROLE PROSODY PLAYS IN DISAMBIGUATION: A STUDY ON MANDARIN

Y.Yang[a,b], S.Gryllia[a], J.S.Doetjes[a], L.L.Cheng[a,b]

[a]Leiden University Centre for Linguistics (LUCL)

[b]Leiden Institute of Brain and Cognition (LIBC)

y.yang@hum.leidenuniv.nl, s.gryllia@hum.leidenuniv.nl, j.doetjes@hum.leidenuniv.nl, L.L.Cheng@hum.leidenuniv.nl

Introduction: In Mandarin, wh-words like *shénme* can receive an interrogative ('what') or a non-interrogative (e.g 'something') interpretation, depending on the preceding discourse and the presence of syntactic licensors. For instance, the string in (1) is ambiguous between a question (1a) and a declarative interpretation (1b).

Aim: Here, we address two research questions: (1) Do speakers use prosody to disambiguate the two interpretations already from the beginning of the sentence? (2) If so, which are the specific cues that they use?

Pre-test: We first ran a pre-test to investigate the availability of the two interpretations. 84 native speakers of Mandarin were asked to read silently sentences like the one in (1) and complete them using a question mark or a full stop. Target sentences (N = 20) were intermingled with 60 fillers. Results showed that 59% of the time participants interpreted the sentences as questions, and 41% of the time as declaratives. On the basis of these results we can conclude that both interpretations are available.

Production experiment: To tackle our research questions we ran a production experiment with a total of 56 target sentences with punctuations (2 clause types × 4 tones on the verb × 7 exemplars of each), intermingled with 132 fillers. Example (2) is an exemplar set. Special attention was given to the tonal composition of the stimuli, which is kept constant across items and conditions; sonorants were used as much as possible. The recordings were made in a sound-proof booth in Tsinghua University in Beijing and the utterances were directly recorded on a computer disk (16 bits, 44100 Hz) using Audacity. The stimuli appeared on a computer screen and participants were asked first to read them silently, and then to utter them. 34 native monolingual speakers of Mandarin from Beijing participated in the experiment.

Prosodic analysis: We analyzed a total of 1904 sentences (56 targets × 34 speakers). We first manually segmented the data at the syllable level, and then extracted duration and F0 measurements using a Praat script (Boersma & Weenink, 2016). Linear mixed effects model was run in R with clause types as fixed-effects factor and item and subject as random factors (Baayen, Davidson & Bates, 2008). **Duration**: In general, questions are shorter than the corresponding declaratives. As shown in Figure 1, the mean duration of subject-, verb plus perfective marker *le-* and *dianr-* constituents are significantly shorter in questions than in declaratives (Subject: Estimate = -0.005, SE = 0.002, t = -2.76, $p < 0.01$, verb plus perfective marker *le*: Estimate = -0.030, SE = 0.001, t = -20.95, $p < 0.01$, *dianr*: Estimate = -0.005, SE = 0.001, t = -5.70, $p < 0.01$). However, the direction changes when looking at the duration of the wh-word; *shénme* in questions is longer than in declaratives (Estimate = 0.012, SE = 0.002, t = 6.99, $p < 0.01$). **F0**: Following Duanmu (2007) and Xu (1997), we extracted $F0_{maximum}$ for T1 (High tone), $F0_{minimum}$ and $F0_{maximum}$ for T2 (Rising tone), $F0_{minimum}$ for T3 (Low tone), and $F0_{maximum}$ and $F0_{minimum}$ for T4 (Falling tone). We converted Hz into semitones (ST); for female speakers ST=12 log2 (Hz/100) and for male speakers ST=12 log2 (Hz/50). Results showed that there were no early F0 differences between questions and declaratives. Questions are higher in register and have a bigger F0 range than declaratives at the verb when it has a T2. In contrast, when the verb has a T4, questions have a smaller F0 range than declaratives. *shénme* in questions has a higher F0 than in declaratives (Estimate = 5.136, SE = 0.131, t = 39.28, $p < 0.01$). This last finding might be related with focus (Hu 2002, Dong 2009).

Conclusion: Our results show that speakers use prosody to disambiguate the two interpretations. When examining duration and F0, it seems that duration is a more prominent

prosodic cue used by speakers already from the utterance onset to disambiguate the two interpretations.

(1)   Táo Wēi zúotiān   ná-le   diǎnr shénme gěi Líu Gāng
       Tao Wei yesterday bring-PERF little what   for Liu Gang
       a. 'What did Tao Wei bring for Liu Gang yesterday?'
       b. 'Tao Wei brought something for Liu Gang yesterday.'
(2)   a. Táo Wēi zúotiān   ná-le   diǎnr shénme gěi Líu Gāng?
          Tao Wei yesterday bring-PERF little what   for Liu Gang
          'What did Tao Wei bring for Liu Gang yesterday?'
       b. Táo Wēi zúotiān   ná-le   diǎnr shénme   gěi Líu Gāng.
          Tao Wei yesterday bring-perf   little something for Liu Gang
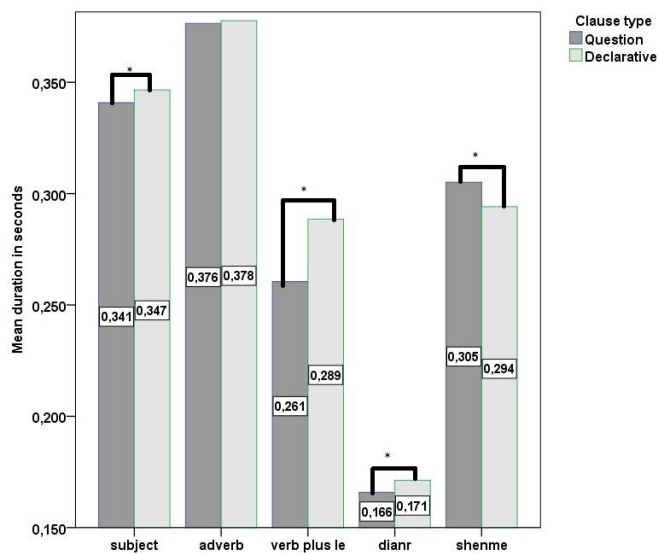   'Tao Wei brought something for Liu Gang yesterday.'

Figure 1. Mean constituent duration across clause type
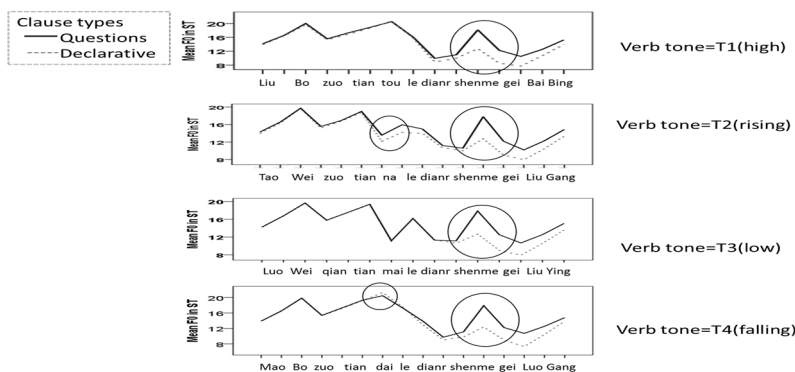

Figure 2. Stylized F0 in ST across clause types and verb tones

REFERENCES: Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory & Language* 59.4 (2008): 390-412. Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.17, retrieved 21 April 2016 from http://www.praat.org/. Duanmu, S. (2007). *The phonology of standard Chinese*. Oxford University Press, USA. Li, Y. H. A. (1992). Indefinite Wh in Mandarin Chinese. *Journal of East Asian Linguistics* 1(2):125-155. Lin, J. W. (1998). On Existential Polarity-WH-Phrases in Chinese. *Journal of East Asian Linguistics* 7(3), 219-255.Xu, Y. (1997). Contextual Tonal Variations in Mandarin. *Journal of Phonetics* 25: 61-83.