

PROSODY ENCODES CLAUSE TYPE ANTICIPATION: EVIDENCE FROM MANDARIN

S.Gryllia^a, Y.Yang^{a,b}, J.S.Doetjes^a, L.L.Cheng^{a,b}

^aLeiden University Centre for Linguistics (LUCL)

^bLeiden 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

Mandarin is a *wh*-in-situ language: *wh*-phrases occupy the same linear position as their non-interrogative counterparts (1a). Comparing the in-situ *wh*-question in (1a) with its declarative counterpart in (1b), we observe that the two clause types are identical up to the *wh*-phrase (*shénme* ‘what’) or the NP (*júzi* ‘oranges’), respectively. Considering examples (1a)-(1b) from a listener’s perspective, the following questions arise. Does the listener need to wait until he hears the *wh*-word or the NP to decide that he is listening to a question or a declarative? Or, alternatively, are there perceptible prosodic differences between the prosody of a *wh*-in-situ question and that of a declarative that appear at the beginning of an utterance so that the listener can anticipate the declarative or question status of the sentence? This paper explores the aforementioned questions by means of an audio sentence completion experiment and claims that the listener can anticipate the difference based on prosodic cues.

EXPERIMENT. On the basis of a previous production study we selected a female native speaker of Beijing Mandarin. Out of her total productions we selected 40 stimuli (20 items × 2 clause types); an acoustic analysis showed that duration is the main factor differentiating questions and declaratives before the *wh*-word/NP. These 40 stimuli were used as a basis for creating three types of audio fragments, each corresponding to a part of the beginning of the sentence. For each type of fragment, one version originated from a question and one from a corresponding declarative. Audio fragment *a* consisted of the subject only. Audio fragment *b* consisted of the subject and the following adverb. Audio fragment *c* consisted of the subject, the following adverb and the verb. Lexical tones were kept constant in the first two audio fragments, while in type *c* the tone of the verb varied (all four tones were included, as to control for the influence of tone). This resulted in a total of 40 audio fragments of each type (120 in total). The audio fragments were presented in three consecutive blocks, corresponding to the fragments of type *a*, *b* and *c* respectively. The experiment was run in MFC Praat (Boersma & Weenink 2013). Participants were instructed to listen to the audio fragment and then complete the sentence choosing one of two possible continuations: a declarative or a question continuation that appeared on screen. 36 participants took part in the experiment at Tsinghua University in Beijing.

RESULTS. In general, listeners were successful in correctly deciding which of the two sentence continuations were intended by the speaker (see Figure 1). More specifically, in audio fragments of type *a* (consisting of only the subject) listeners chose a question continuation for fragments originating from questions in 59.6% of the times. For fragments originating from declaratives, they chose a declarative continuation in 64.6% of the times. There was a significant association between the clause type intended by the speaker and the listeners’ responses, $\chi^2 = 84.31(1)$, $p < .001$. For audio fragment *b* the percentages were 63.9% and 72.2%, for questions and declaratives respectively. There was, again, a significant association between the intended by the speaker clause type and the listeners’ responses, $\chi^2 = 189.09(1)$, $p < .001$. In audio fragment *c*, listeners chose a question continuation in 56.5% of the times when the speaker intended a question, while they chose a declarative continuation in 78.1% of the times when the speaker intended a declarative. Again, there was a significant association between the intended clause type and the listeners’ responses: $\chi^2 = 180.59(1)$, $p < .001$. Besides this, we analysed the data per audio fragment modelling continuation type likelihood using logit mixed-effect models. All the analyses were run in R using the lme4 software package. We started with a null model which included continuation type as a dependent variable and participants and items as random factors. We

then ran another model adding the clause type intended by the speaker to see whether the model improved. Our results show that the intended clause type is a good predictor of the listeners' responses. On the basis of these results, we can conclude that listeners of questions and corresponding declaratives can make use of prosody to anticipate the clause type.

Examples

(1) a. Bái Wēi zúotiān bāo-le shénme gěi Lúo Yīng? [wh in situ question]
 Bai Wei yesterday peel-PERF what to/for Luo Ying
 ‘What did Bai Wei peel for Luo Ying yesterday?’

b. Bái Wēi zúotiān bāo-le júzi gěi Lúo Yīng. [declarative]
 Bai Wei yesterday peel-PERF oranges to/for Luo Ying
 “Bai Wei peeled oranges for Luo Ying.”

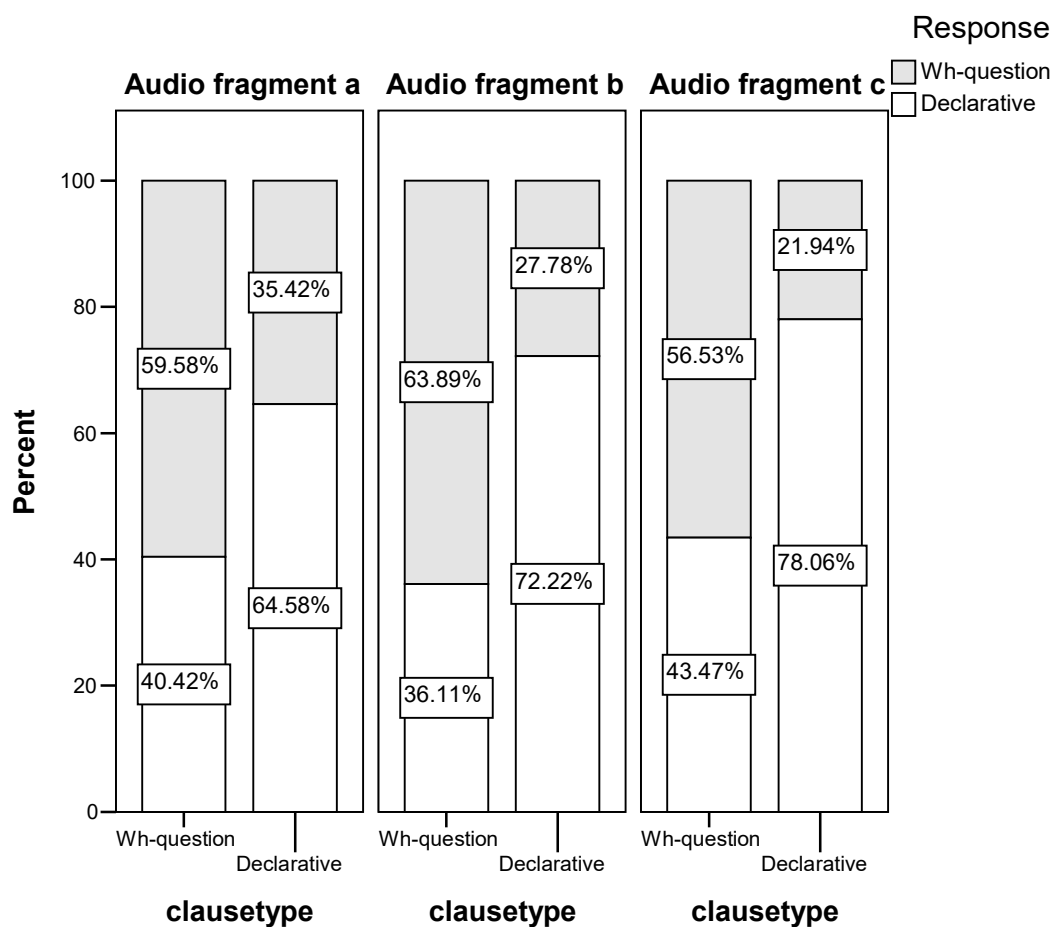


Figure 1. Listeners' responses to fragments originating from questions or declaratives across three types of audio fragments

References

Bates D., M. Maechler, B. Dai (2008) lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-28.
 Boersma, P., D. Weenink (2013) Praat: doing phonetics by computer [Computer program]. Version 5.3.51, retrieved 2 June 2013 from <http://www.praat.org/>