

## Do listeners detect prediction errors in lexical tone? An investigation with tone sandhi in Mandarin Chinese shows unexpected prediction inhibition

Yiling Huo<sup>a</sup>, Suiping Wang<sup>b</sup>, and Wing-Yee Chow<sup>a</sup>

<sup>a</sup>Division of Psychology and Language Sciences, University College London, London, UK; <sup>b</sup>Philosophy and Social Science Laboratory of Reading and Development in Children and Adolescents (South China Normal University), Ministry of Education, Guangzhou, China

The ability to predict is key to real-time language comprehension. Previous research found that humans can not only use a wide array of information to predict, but they are also very good at detecting prediction errors and revising their predictions on the fly [1]. A recent eye-tracking study showed that native Mandarin Chinese speakers use their knowledge about lexical tones and a specific tone sandhi rule (T3 sandhi) to predict upcoming language [2]. In this study, we used a visual-world eye-tracking paradigm to ask whether listeners can detect an early sign of prediction error purely in the lexical tone of a numeral and use it to revise their noun prediction accordingly. Results suggest that **listeners can use prediction-inconsistent classifiers, but not tone, to revise their noun predictions**. However, an unexpected 'resetting' eye-movement pattern observed prior to the numeral suggests that listeners may have inhibited their predictions to some degree.

**Method.** In the present experiment, participants ( $n=38$ ) saw pairs of objects on the screen while listening to highly constraining sentences in Mandarin Chinese and were asked to select the object mentioned in the sentence as quickly as possible. Feedback was given on accuracy and speed after each trial with a traffic light feedback system (2). While 2/3 of all sentences in the experiment (100 fillers) ended with the expected noun, the other 1/3 (10 fillers and 40 experimental items) **ended with an unexpected but congruous noun**. In the experimental items, the target noun always appeared in a numeral-classifier-noun phrase (critical NP; 1); it was always depicted alongside the expected noun on the screen, and the nouns were always associated with distinct nominal classifiers. **In the Different Tone condition, the nouns were chosen such that the numeral in the critical NP had a different lexical tone (as a result of undergoing tone sandhi) depending on the noun and the classifier it was associated with. Thus, prior to the classifier and noun, the tone of the numeral itself could already signal to the participants that their noun prediction was wrong.** Meanwhile, in the Same Tone condition, the two nouns were chosen such that the numeral always had the same lexical tone and as such was uninformative. Further, half of the experimental items had the numeral *yi* ('one') that undergoes the *yi* sandhi, while the other half had the numeral *liang* ('two'), where the T3 sandhi is applicable.

**Results and Discussion.** Divergence point analysis using bootstrapping methods [3] did not reveal a significant difference in the onset of looks to the unexpected target between the Same Tone and Different Tone conditions. Mixed-effects linear model did not reveal significant differences between conditions in log reaction times, either. Thus, we did not find evidence for listeners' ability to detect prediction error in the numeral's tone alone. However, we observed a somewhat unexpected eye movement pattern around the onset of the critical NP (Figs 1, 2). Although listeners were slightly more likely to look towards the expected object during the sentence context, this preference seemed to have 'reset' by the start of the critical NP, such that looks to both objects were at chance level until the classifier is heard. **This suggests that participants may have strategically inhibited their noun prediction computations in anticipation of the critical NP.** Under time pressure posed by the choosing-while-listening task, listeners may have chosen to wait and rely on the classifier to predict the noun, as it was perceptually salient and highly reliable. Future studies will be needed to investigate this possibility. The present results may still be taken to suggest that listeners cannot use the tone of the numeral itself to identify the upcoming noun, but we cannot rule out the possibility that listeners' ability to detect a prediction-

mismatch tone in the numeral may have been hampered when they chose to inhibit or 'reset' their prediction.

### References

- [1] Chow, W. Y., & Chen, D. (2020). Predicting (in) correctly: listeners rapidly use unexpected information to revise their predictions. *Language, cognition and neuroscience*, 35(9), 1149-1161.
- [2] Chow, W.Y., & Huo, Y. (2020) Listeners can use tone information to predict: Evidence from Mandarin Chinese. *Poster at the 33rd annual CUNY Human Sentence Processing Conference*.
- [3] Stone, K., Lago, S., & Schad, D. J. (2021). Divergence point analyses of visual world data: Applications to bilingual research. *Bilingualism: Language and Cognition*, 24(5), 833-841.


#### (1) Experiment design.

Lexical tones relevant to the experimental manipulation are spelt using numbers.

(2) The traffic light feedback system. Time of response is judged against the noun onset.

Different Tone

“安妮在星巴克买了一块蛋糕”  
 Anni zai xingbake mai-le **yi2** kuai4 dangao  
 Anne at Starbucks buy-ASP one piece cake  
 “At Starbucks, Anne bought a piece of cake.”

  
 一杯咖啡  
 yi4 bei1 ka1fei1  
 one cup coffee  
 'a cup of coffee'

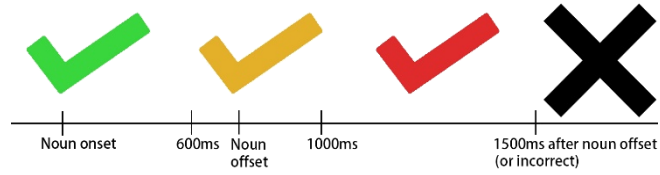
  
 一块蛋糕  
 yi2 kuai4 dan4gao1  
 one piece cake  
 'a piece of cake'

Same Tone

“安妮在星巴克买了一瓶水”  
 Anni zai xingbake mai-le **yi4** ping2 shui.  
 Anne at Starbucks buy-ASP one bottle water  
 “At Starbucks, Anne bought a bottle of water.”

  
 一杯咖啡  
 yi4 bei1 ka1fei1  
 one cup coffee  
 'a cup of coffee'

  
 一瓶水  
 yi4 ping2 shui3  
 one bottle water  
 'a bottle of water'



#### (3) Cloze probability profile of experimental items

Cloze prob. of noun	Mean	Min	Max	Median
Expected	0.67	0.45	0.97	0.66
Unexpected	0.03	0.00	0.26	0.00

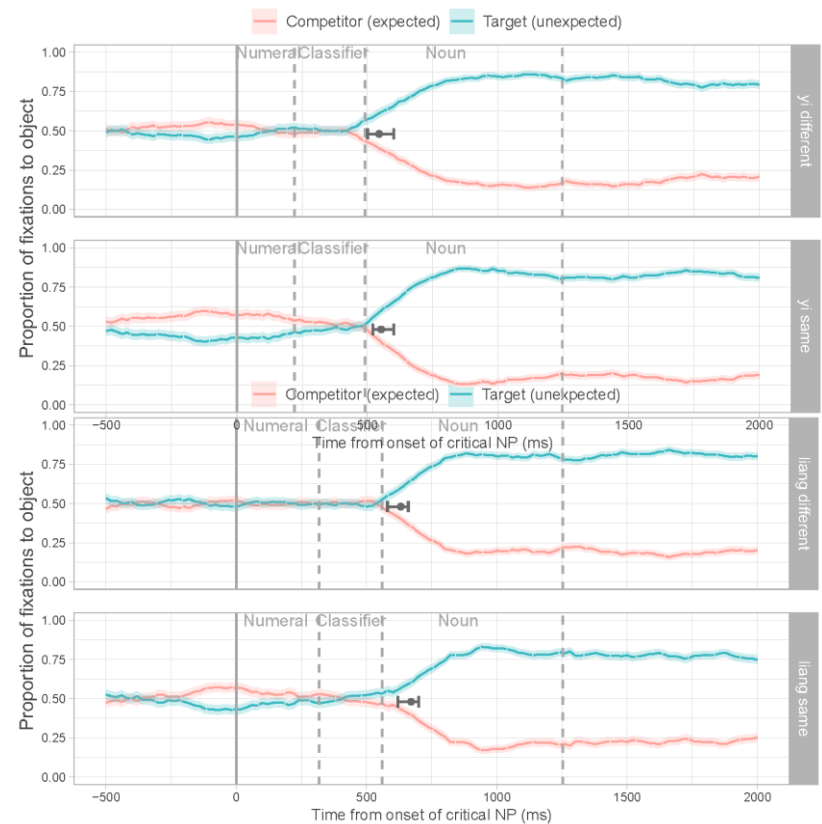


Figure 1. Proportion of fixations to each object with sentences containing numeral yi ('one'), which undergoes the yi sandhi. Black dots and horizontal error bars indicate the mean onset of looks to the target object and 95% credible intervals. Divergence point analysis did not show significant difference in onset of fixations to target between conditions.

Figure 2. Proportion of fixations on each object. Sentences containing the numeral liang ('two'), which undergoes the T3 sandhi. Black dots and horizontal error bars indicate the mean onset of looks to the target object and 95% credible intervals. Divergence point analysis did not show significant difference in onset of fixations to target between conditions.