The sound of syntax: Prosodic disambiguation for syntactic branching in a tone language
Zeping Liu and Chien-Jer Charles Lin (Indiana University Bloomington)

Introduction: Prosodic information is crucial in resolving syntactic ambiguities. Previous studies have suggested that phrasal boundaries and stress patterns influence structural analysis and interpretation during spoken language processing [1]. However, due to the absence of a consistent one-to-one prosody-syntax mapping and arbitrariness of this mapping across languages, prosodic disambiguation remains language-specific and challenging, particularly for Mandarin, where pitch is primarily used for lexical tones. Given these, this study aims to explore how Mandarin native speakers prosodically resolve structural ambiguity, focusing on an ambiguous phrase *Adj N1 N2* in Mandarin. This phrase is structural ambiguous between a left-branching (LB) and a right-branching (RB) analysis, involving whether the adjective modifies N1 or N2. For example, *fense/xingxing/maozi* (*pink star hat*) can refer to either "a hat with pink stars" (LB) or "a pink hat with stars" (RB). Previous work on Japanese and Korean suggests a syntactic preference for LB though disambiguating prosody differs in these languages (Japanese: pitch change; Korean: accentual phrasing) [2-3]. It remains unclear whether the LB preference is universal across languages, shaped by incremental and simplicity principles in human parsing (**Exp. 1**). Additionally, debate exists on whether prosody is context-dependent, used only in situations requiring disambiguation [4-5]. Given these, **Exp. 2** examines whether and how Mandarin speakers prosodically differentiate the two analyses across different contexts.

Materials & Design: Twenty-three ambiguous phrases were used, each following a "color+pattern+object" sequence all with disyllabic words. **Exp. 1** used a phrase-picture matching task where participants (*N*=37) determined whether acoustically neutral ambiguous phrases matched the pictures (**Fig. 1**, left panel). Responses and reaction times (RTs) were recorded. **Exp. 2** used a sentence reading-out task with an audience design [6-7] to investigate how speakers use prosody to disambiguate the phrase in ambiguous versus unambiguous context. In this task, 16 Mandarin native speakers were presented with visual displays, where branching type (LB or RB) and Contrastiveness (No Contrast vs. Contrast, depending on whether the visual scene included only LB or RB images or both) were manipulated across 4 lists (**Fig. 1**, right panel). Participants were given a written sentence (e.g., *Please click on the pink star hat on the screen*) referring to the selected image, and they needed to read the sentence out. Participants were informed that their production would be recorded to help a hypothetical listener select the same image. The analysis focuses on acoustic correlates (e.g., timing variations, pitch patterns, and intensity) associated with branching disambiguation, since the two interpretations should have different prosodic structures according to the prosody-syntax interface.

Results & Discussion: In the matching responses, RTs for LB images were faster than RB images (*Estimate* = 46.26, *t* = 2.184, *p* < .05), supporting the preference for the LB analysis and aligning with the local parsing constraint in incremental sentence processing. Phonological analyses of speakers' production in Exp. 2 revealed that silent pauses were more frequently inserted between Adj and N1 in the Contrast RB condition (31.5%) than in the No Contrast RB condition (8.7%). Combined with the longer RB pause and Adj lengthening in the Contrast RB condition (**Fig. 2**), this suggests a significant prosodic boundary between Adj and N1 when the RB analysis is intended. The contrast effect was also observed in the LB condition, where a pause was more likely to be placed between N1 and N2 when the visual display was ambiguous. These findings support the context-dependent nature of prosodic disambiguation [4], providing further evidence that additional prosodic marking is needed for less preferred structure (i.e., RB analysis) in disambiguation. No significant differences were observed in pitch patterns or intensity across conditions (**Fig. 3**). Taken together, timing variations, mainly silent pause and word lengthening, serves as the major prosodic realization of branching disambiguation in Mandarin—a tone language where pitch-based prosodic cues are primarily lexical and not reliably intonational [8].
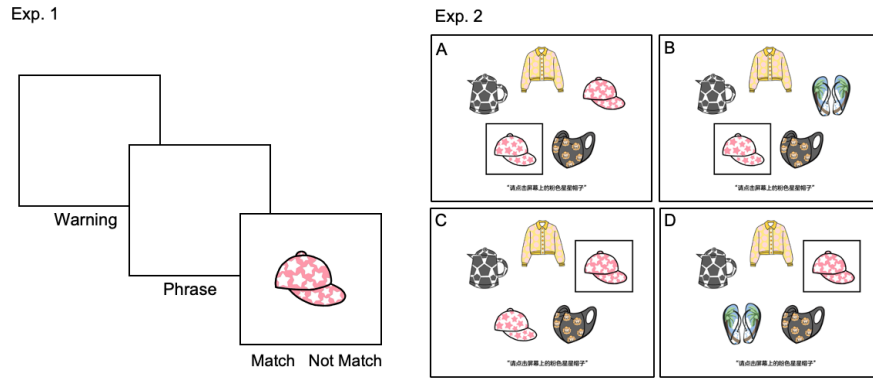
Fig. 1 Design in Exp.1 & 2. In Exp. 2, selected images are Contrast LB (A), NoContrast LB (B), Contrast RB (C), and NoContrast RB (D). Below images is the written sentence to be read out by the speakers.
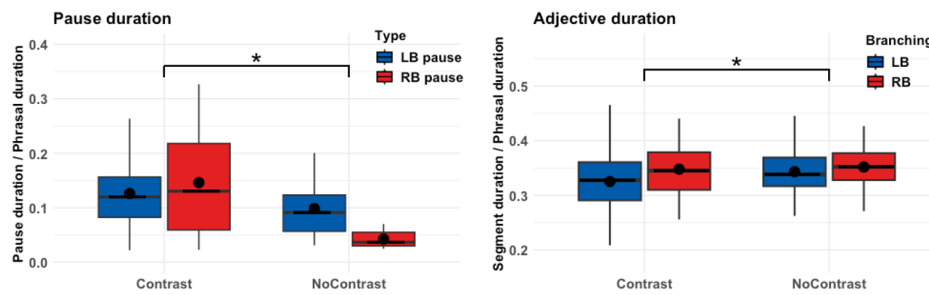

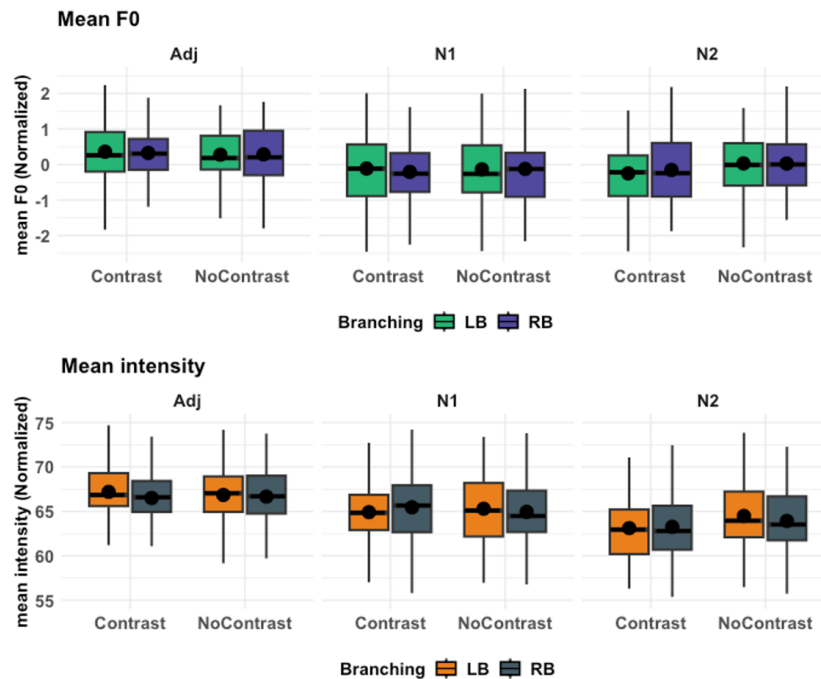**Fig. 2** Pause and adjective duration for each condition.

Fig. 3 Mean F0 and intensity for each segment within the phrase in each condition.

References: [1] Warren, 1999. In S. Garrod & M. J. Pickering (eds.), *Language Processing*. [2] Hirose, 2020. *Language and Speech*. [3] Schafer & Jun, 2002. In Nakayama (eds.), *Sentence Processing in East Asian Languages*. [4] Snedeker & Trueswell, 2003. *Journal of Memory and Language*. [5] Speer, 2011. *Laboratory Phonology*. [6] Lockridge & Brennan, 2002. *Psychonomic bulletin & review*. [7] Clark & Carlson, 1982. *Language*. [8] Ouyang & Kaiser, 2015. *Language, Cognition, and Neuroscience.*