

Effects of visual and discourse contexts and prosody on referential resolution

Kiwako Ito (The Ohio State University), Chie Nakamura (Keio University), & Reiko Mazuka (RIKEN BSI)
ito@ling.osu.edu

Contrastive prosody in speech comprehension; Visual world paradigm; Japanese

Successful discourse comprehension requires constant updates of what has been discussed and reasonable anticipations for what is likely to be discussed next. Previous studies demonstrate the immediate impact of contrast-evoking prosody on the referential resolution [1, 2, 3], yet the relationship between the context-driven referential salience and the role of prosody in guiding referential resolution has not been well investigated. The present study examines how contrast-evoking prosodies in Japanese is interpreted when the visual and discourse contexts elicit changes in salience among candidates for the subsequent referential resolution.

Experimental slides presented eight animals, each holding an object. Critical slides included a trio (e.g., three gorillas holding a racket, an umbrella and a ball, respectively) a twin with two objects repeated from the trio (two raccoons with a racket and an umbrella), a singleton (a rabbit with an umbrella) and a distracter twin (two frogs with a balloon and a baseball bat). Participants followed sequential instructions (Example 1). Q1 had no particular prominence. Q2 sentences were recorded with (1) prominence on the object noun, (2) pre-focal attenuation on the object noun and prominence on the animal noun, (3) no particular prominence. If the prominence on the object evokes contrast with Q1, (1) should lead to higher fixations to the contrastive competitor (e.g., gorilla with umbrella) than (2) and (3). If the pre-focal attenuation evokes contrast on the upcoming animal, (2) should lead to higher fixations to the animal other than the competitor than (1) and (3). The mention of a trio member in Q1 was predicted to elicit three levels of salience among the Q2 candidates holding the same object (umbrella): the trio member (e.g., gorilla) should become most salient due to the explicit mention; the twin whose counterpart has the same object as in Q1 (e.g., raccoon) should be moderately highlighted; the unmentioned singleton (e.g., rabbit) should be least salient.

The eye-movement patterns revealed interesting discourse-driven attention shift and the salience-driven use of prosodic cues. During the 1s exposure to the slide before linguistic input, participants fixated the distracter twin most often and the singleton least. Then, Q1 led to immediate increase in the fixations to trio members. Upon hearing the Q2 object (e.g., umbrella), fixations to the trio competitor (gorilla with umbrella) were higher than to the Q2 target (raccoon) regardless of the prosody. However, this anticipation was stronger for (1) and (2) than for (3), suggesting that both prominence on the object and pre-focal attenuation enhanced the context-driven contrastive interpretation. Interestingly, (2) led to higher fixations to the singleton (rabbit) than (1) and (3), indicating that pre-focal attenuation directed attention to previously least attended candidate. In sum, the results suggest that the discourse context overwrites the visual-based initial referential anticipation. Prosodic cues may not only facilitate the discourse-driven anticipation but also evoke alternative referential interpretation. The time courses of fixation patterns suggest that these alternative interpretations were simultaneously available, reflecting the parallel evaluations of candidates for referential resolution.

Example 1: Q1: *raketto-o motteru gorira-wa doko?*
racket-ACC holding gorilla-TOP where
'Where is the racket-holding gorilla?'

Q2: (1) *Jaa, KASA-o motteru tanuki-wa doko?*
(2) *Jaa, kasa-o motteru TANUKI-wa doko?*
(3) *Jaa, kasa-o motteru tanuki-wa doko?*
then umbrella-ACC holding raccoon-TOP where
'Then, where is the umbrella-holding raccoon?'

References

- [1] Dahan, D., Tanenhaus, M.K., & Chambers, C.G. (2002). Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, 47, 292-314.
- [2] Weber, A., Braun, B., Crocker, M. W. (2006). Finding Referents in Time: Eye-Tracking Evidence for the Role of Contrastive Accents. *Language and Speech*, 49 (3), 367-392.
- [3] Ito, K. & Speer, S. R. (2008). Anticipatory effect of intonation: Eye movements during instructed visual search. *Journal of Memory and Language*, 58, 541-573.