

The modulation of predictive processing in the context of multitalker babble

Carrie N. Jackson, Theres Grüter, Navin Viswanathan

Research on predictive processing shows that language users modulate their predictions based on top-down assumptions regarding the reliability of bottom-up linguistic cues and the utility of using those cues to predict upcoming input (Hopp, 2016; Kuperberg & Jaeger, 2016; Ryskin & Nieuwland, 2023). Recent studies have shown that predictive processing is delayed when native (L1) listeners listen to sentences produced with L2-accented speech (Porretta et al., 2020) or sentences that contain multitalker babble (Carroll et al., 2022), due to the increased effort needed to process the incoming speech stream and increased uncertainty about the reliability of the linguistic input. However, most research to date has considered such modulations in prediction when they stem from changes to the critical bottom-up linguistic cues themselves. Less research has considered how top-down assumptions influenced by changes in the general reliability of the signal in the broader speech context may modulate listeners' reliance on bottom-up linguistic cues to predict upcoming input, even when the critical cues themselves remain reliable. To address this question, the present study investigates whether the presence of multitalker babble, even when only occurring in a limited number of filler sentences, is sufficient to modulate whether speakers engage in semantic prediction while listening to target sentences presented with a silent background.

L1 English listeners ($n=56$, ongoing) listened to English sentences while looking at a display of four images (Fig. 1). Each visual display contained images of the target noun (e.g., ladder), two competitor nouns that were semantically related to the sentence subject (e.g., axe, hose) and an unrelated distractor noun (e.g., paperclip; materials adapted from Porretta et al., 2020). We tracked participants' eye-movements to measure whether they used semantic information from a lexical verb that restricted the choice of direct object (e.g., *climb*, constraining verb; 1a) to anticipate the upcoming target noun compared to a lexical verb that could take any number of objects (e.g., *need*, neutral verb; 1b). For participants in the "silent background" group ($n=29$), all experimental and filler sentences were spoken by a single male speaker in quiet. For participants in the "babble background" group ($n=27$), all experimental sentences were identical to those in the silent background group, but 50% of filler items contained competing two-talker female English babble at an SNR of -10 dB. This babble had a variable onset such that it always started after the subject but before the end of the lexical verb in the target sentence.

1a. The fireman will climb the ladder. (constraining verb)

1b. The fireman will need the ladder. (neutral verb)

Divergence point analyses (Stone et al., 2021) revealed earlier looks to the target for constraining- versus neutral-verb sentences, based on (nearly) non-overlapping bootstrapped CIs, in both the silent and babble groups (Fig. 2). This provides preliminary evidence that participants in both groups used verb semantics to anticipate the target noun. However, (nearly) non-overlapping bootstrapped CIs for constraining-verb sentences in the silent versus babble background groups preliminarily suggest that this predictive processing was delayed in the babble background group. In contrast, the two groups exhibited no difference in divergence points for neutral-verb sentences, based on overlapping bootstrapped CIs.

In sum, listeners in both groups engaged in semantic prediction but the presence of multitalker babble in the broader listening context delayed these effects in the babble group, even though the experimental sentences themselves were spoken by a single speaker under silent background conditions and, as such, semantic prediction remained equally facilitatory for both groups. This adds to evidence that listeners use top-down information based on properties of the broader listening context to adjust expectations regarding the reliability of the input and, in turn, the utility of engaging in prediction during language processing.

Figures

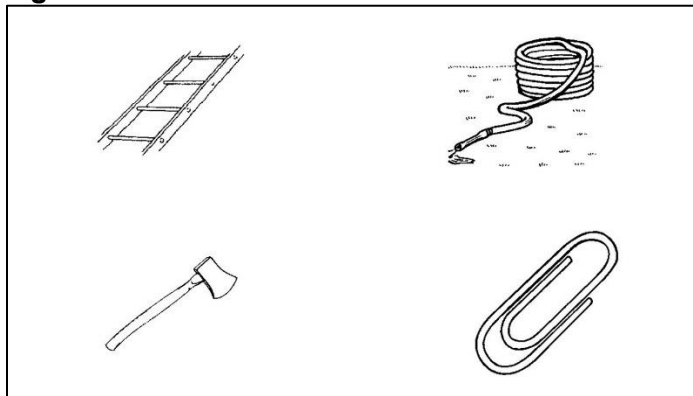


Figure 1. Sample image display (for sentences 1a. and 1b.)

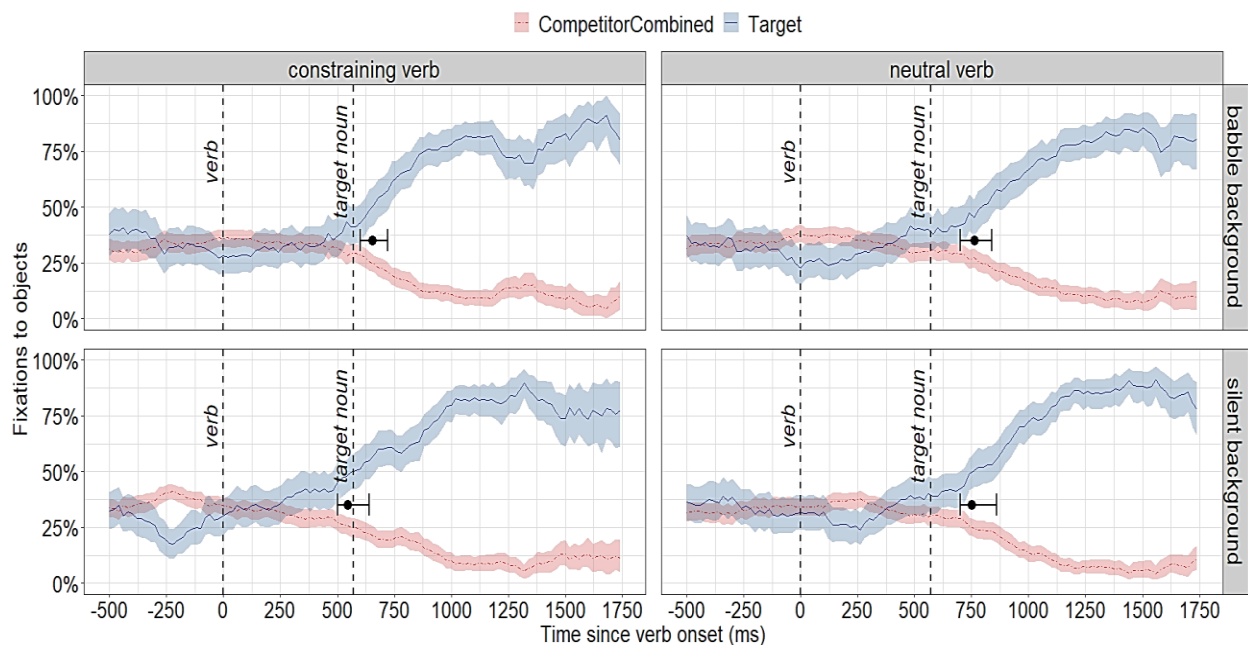


Figure 2. L1 speakers' fixations to target noun in constraining- and neutral-verb sentences for babble and silent background groups. Divergence points (with bootstrapped 95% confidence intervals) in black.

References

- Carroll, R., Patarroyo, A.G., & Hopp, H. (2022). Predictive L2 sentence processing in noise: Differential effects across linguistic cues. Paper presentation at the *International Symposium on Bilingual and L2 Processing in Adults and Children*, Tromsø, Norway, August 4-5.
- Hopp, H. (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, 32(2), 277-307.
- Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, 31(1), 32-59.
- Porretta, V., Buchanan, L., & Järvikivi, J. (2020). When processing costs impact predictive processing: The case of foreign-accented speech and accent experience. *Attention, Perception, & Psychophysics*, 82(4), 1558-1565.
- Ryskin, R., & Nieuwland, M. (2023). Prediction during language comprehension: what is next? *TICS*, 27(11), 1032-1052. <https://doi.org/10.1016/j.tics.2023.08.003>
- 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.