**Re-assessing the speech-to-eye lag in language comprehension**

Jéssica Gomes[1], João Veríssimo[1], Dan Parker[2], Sol Lago[3]
[1]University of Lisbon; [2]Ohio State University; [3]Goethe University Frankfurt

Visual world research on language comprehension typically assumes that—in the absence of coarticulation—it takes about 200 ms for information in the speech signal to affect eye movements. This assumption is crucial for estimating psycholinguistic effects like prediction in visual world studies with diverse populations, including child and adult native speakers [1–3], second language learners [4], and older adults [5]. However, studies that experimentally established the 200 ms speech-to-eye lag were performed on relatively homogeneous samples of English-speaking university students [6,7] and used a saccade-detection method with potential issues, including the removal of large numbers of trials (more than 50% in [7]) and the use of statistical models that did not account for between-participant variability. To address these issues, we performed a conceptual replication of [7] using similar materials but a more diverse population: second language (L2) learners of English. We diagnosed the speech-to-eye lag in L2 learners using both the saccade-detection method in [7] and a novel method based on generalized additive models [8], which allowed estimating the speech-to-eye lag for each participant and the role of individual predictors, namely age of L2 acquisition and L2 proficiency.

**Method.** Seventy-nine German learners of English (mean age of English acquisition = 7.7 years, mean score on the English proficiency test LexTALE = 84%) performed a visual world eye-tracking experiment. In each trial, participants heard instructions like “Click on the book, please!” while looking at displays with four objects: the target object, an associated distractor and two other objects (Figure 1A). The speech-to-eye lag was defined as the earliest time at which the model-estimated proportion of fixations to the target was significantly greater than to the associated distractor [8]. Participants also completed a demographic questionnaire and the English proficiency test LexTALE [9]. We used these measures to assess the effects of age of acquisition and L2 proficiency.

**Results and discussion**. At the group level, English learners showed a slightly longer mean speech-to-eye lag than the 200 ms lag previously reported for native speakers: 261 ms, 95% CI [175, 319] ms. While the effect’s confidence interval included [7]’s original 200 ms estimate, the later mean latency of our speech-to-eye lag supports previous evidence that lexical retrieval—a cognitive process contributing to the lag—is slower in second than first-language speakers [10]. The saccade-detection method of [7] yielded a similar estimate of the group-level lag (260 ms), providing an internal validation of the novel method. There was evidence of between-participant variability (Figure 2A). Additionally, we found an effect of age of acquisition: Participants who began acquiring English earlier showed shorter lags (Figure 2B). There was no evidence of an effect of L2 proficiency as measured by the LexTALE, likely due to the high proficiency of our sample (Figure 2C). Overall, these results have useful implications for visual world research because they suggest that the originally established 200 ms lag should not be taken as an universal estimate but rather experimentally evaluated in the population of interest, also taking into account the potential role of individual differences [11,12]**.**

**Figure 1.** (A) Sample visual display in the experiment and (B) Percentage of looks to the objects across participants. Following [7], one of the three distractors was named “associated distractor” and used as a baseline in the saccade-detection method to compare with the target object. The determiner and noun were cross-spliced to eliminate co-articulatory cues at the determiner.

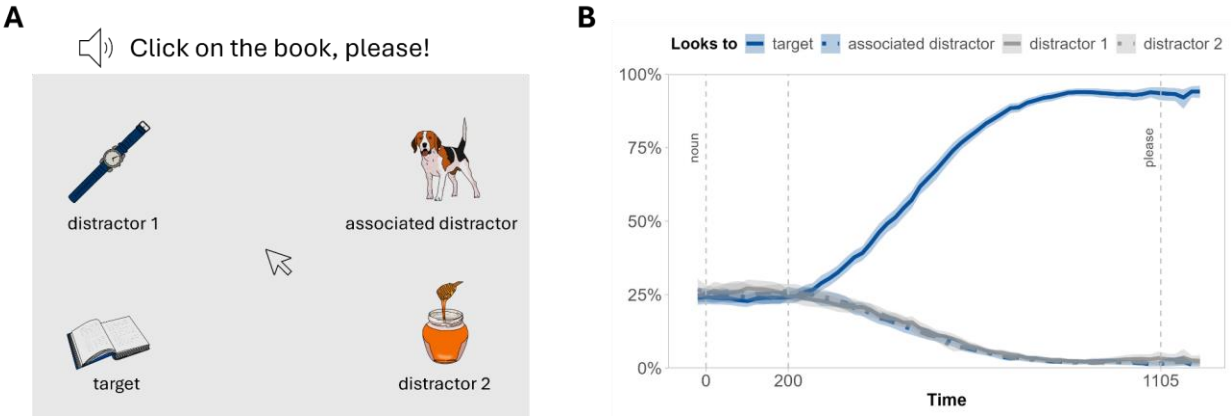
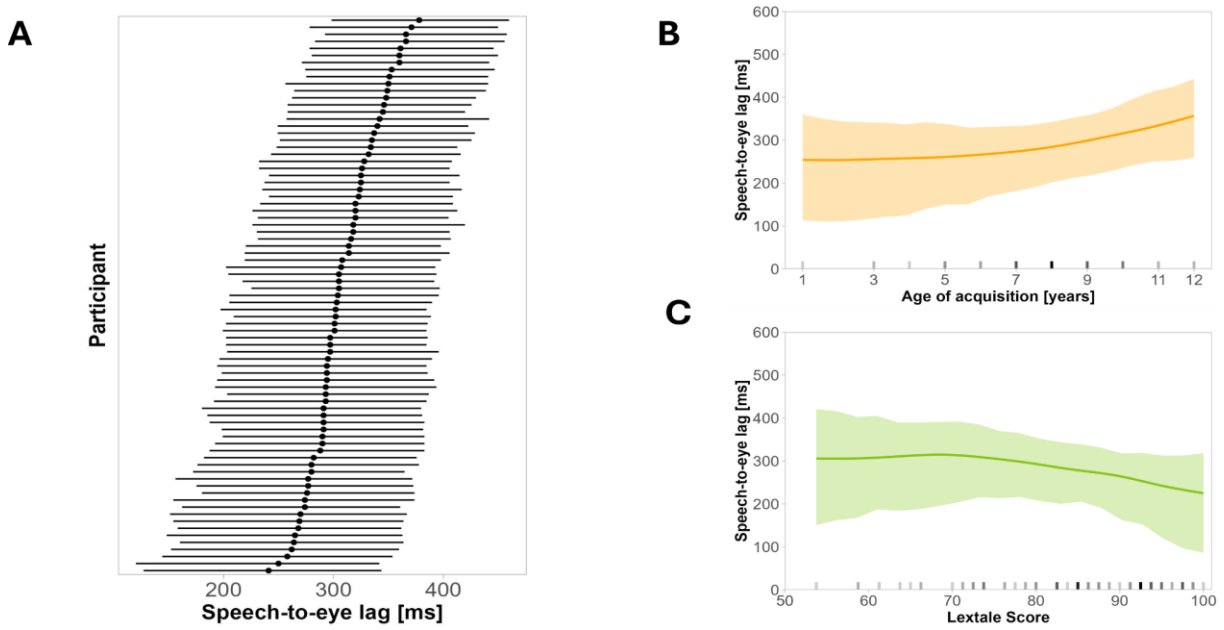



**Figure 2.** (A) Size of the speech-to-eye lag for each participant. (B) Age of acquisition effect and (C) LexTALE effect. The horizontal lines (A) and ribbons (B–C) show 95% confidence intervals. The number of participants per age of acquisition or LexTALE value is shown by a rug on the x-axis.




**References:** [1] Matin et al. (1993). *Percept. Psychophys.* [2] Milburn et al. (2016). *Lang. Cogn. Neurosci*. [3] Järvikivi et al. (2014) *Lang. Cogn. Neurosci*. [4] Ito et al. (2017) *JML*. [5] Noriet et al. (2016). *Aging Neuropsychol. Cogn*. [6] Tanenhaus et al. (1995). *Science.* [7] Salverda et al. (2014). *JML*. [8] Veríssimo & Lago (2023). *AMLaP 2023.* [9] Lemhöfer & Broersma (2012). *Behav. Res. Methods.* [10] Sullivan et al. (2018). *Biling*. [11] Kidd et al. (2018). *Trends in Cog Sciences*. [12] Cunnings & Fujita (2021). *Appl. Psycholinguist.*