

# Lexical Access during Naturalistic Listening in L1 and L2 Speakers of English

Tanya Levvari, Jarom Larman, Briony Waite, & Jesse Snedeker  
Harvard University

**Introduction.** In a natural, spoken conversation we hear two or more words per second. Even at this rapid rate, native, adult speakers predict words before encountering them. <sup>e.g.1</sup> Lexical prediction is a top-down process in which information from context is used to pre-activate possible upcoming word meanings. <sup>1,2</sup> The degree to which L2 learners rely on top-down cues from context during lexical processing has been debated. A recent study finds that proficient L2 speakers show similar use of top-down cues as L1s, generating predictions that modulate the N400 response. <sup>3</sup> However, this study presents words in isolation, while earlier studies using more complex input find that L2 speakers show predictive deficits. <sup>4,5</sup> In the current study, we evaluate lexical processing in monolingual L1 and bilingual L2 English speakers as they listen to a story in order to evaluate reliance on top-down prediction during a naturalistic listening task. We are specifically interested in seeing the degree to which the N400 response in L1 and L2 English speakers reflects bottom-up lexical features (e.g. Frequency) vs. top-down prediction (Cloze Probability).

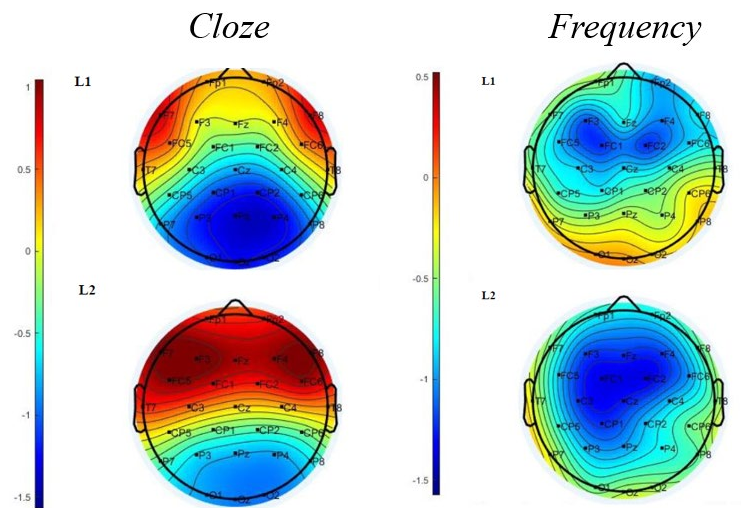
**Methods.** The current sample currently includes 14 L1 English participants (mean age=21.5, mean PPVT = 111.4) and 23 L2 English participants (mean age = 20.5, mean PPVT = 94.1). L2s began learning English after the age of 5, vary in their L1, and are fluent in English. However, their PPVTs are significantly lower than L1s ( $p < 0.05$ ). We use the naturalistic Storytime Paradigm: participants listen to a 30-minute children's story in English while we recorded continuous EEG. Our analyses focus on the amplitude of the N400, an ERP component related to lexical access. <sup>6</sup> We coded all 2,298 content words in the story for Frequency (SUBTLEXus<sup>7</sup>), their Association to preceding content word (Word2Vec), and their Predictability (cloze probability given the context based on a norming sample of L1 speakers). In addition, we control for factors including concreteness<sup>8</sup>, acoustic length, and serial position in the story and baseline effects of Frequency, Association, and Predictability of the immediately preceding and subsequent words.

**Results.** To evaluate the effects of Frequency, Association, and Cloze on the N400 response, we looked at whether these significantly predicted the N400 above and beyond any effects of control variables. Data collection is on-going; however, preliminary results show that Frequency significantly predicts N400 size when both L1 and L2 participants are included in the analyses ( $\beta=0.20$ ,  $SE=0.1$ ,  $p<0.05$ ). Cloze also significantly predicts the N400 response above and beyond the effects of Frequency and the other control factors ( $\beta=0.29$ ,  $SE=0.09$ ,  $p<0.01$ ). Semantic Association does not predict N400 responses. There is however a marginal interaction of Group and Frequency ( $\beta=-0.13$ ,  $SE=0.07$ ,  $p=0.08$ ). Looking separately at each group, we see that L1s show only a significant effect of Cloze ( $\beta=0.33$ ,  $SE=0.14$ ,  $p<0.05$ ), with no effects of Frequency or Semantic Association. However, unlike our L1s, the L2 participants show a highly robust effect of Frequency ( $\beta=0.31$ ,  $SE=0.12$ ,  $p<0.01$ ). Our L2 group also showed a significant effect of Cloze ( $\beta = 0.24$ ,  $SE = 0.10$ ,  $p = 0.02$ ). However, visual analyses of the scalp maps (See Figure 1) suggest different scalp distributions across our two participant groups.

**Discussion.** Our results thus far suggest that L1 and L2 speakers of English differ in their reliance on bottom-up vs. top-down cues during online lexical access. Specifically, during naturalistic listening, the amplitude of the N400 reflects Word Frequency only in the L2s. On the other hand, both L1 and L2 English speakers rely on top-down cues from context. These findings support prior work showing differences in L2 online language comprehension and extend these findings to a naturalistic setting. However, planned analyses will further investigate differences in scalp distribution and possible effects of fluency in our full sample.

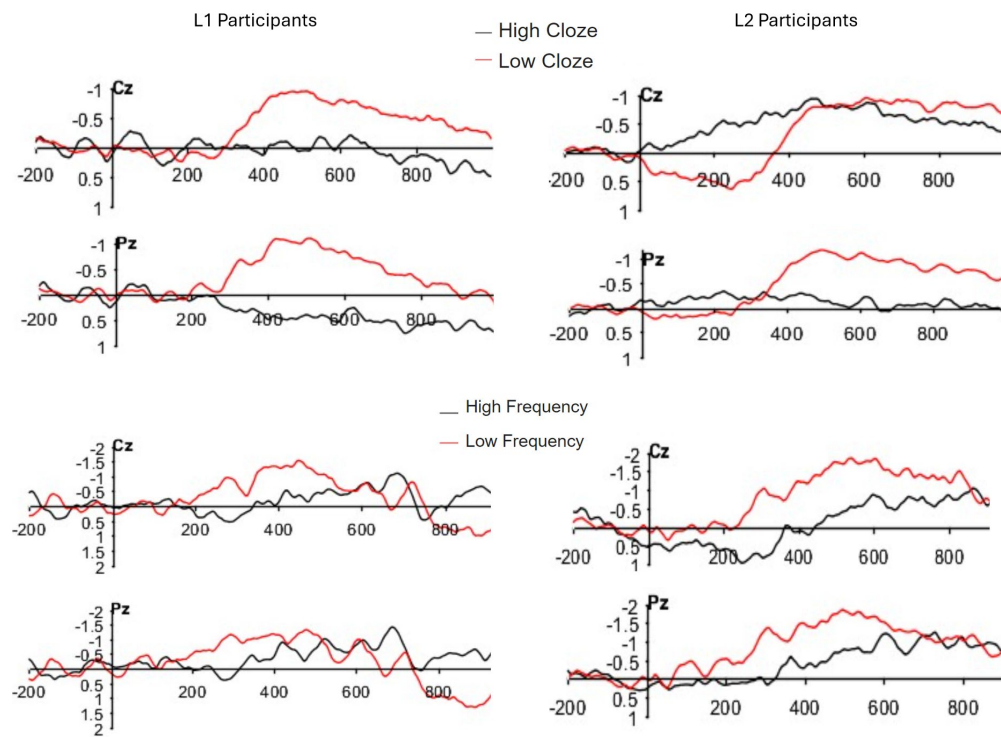
**References.** <sup>1</sup>Kuperberg & Jaeger, 2016; <sup>2</sup>Altmann & Mirković, 2009; <sup>3</sup>Banon & Martin, 2023; <sup>4</sup>Foucort et al., 2014; <sup>5</sup>Martin et al., 2013; <sup>6</sup>Kutas & Federmeier, 2011; <sup>7</sup>Brysbaert & New, 2009; <sup>8</sup>Mikolov et al., 2008, <sup>9</sup>Balota et al., 2007

Figure 1:



Topographic maps of amplitude differences between Low and High Cloze (Left) and Frequency (Right) words in L1 (top) and L2 (bottom) participants. Our L2 participants display a robust negativity to Low Frequency words.

Figure 2:



ERP responses to Low and High Predictability words in L1 and L2 (Top) and Low and High Frequency words (Bottom) at electrode centro-parietal sites (Cz and Pz)