Revising the linking hypothesis between reading times and processing difficulty using a computational model of reading.

An important goal of sentence processing studies in psycholinguistics is to account for how, by combining word-level information, we obtain a sentence-level representation. This includes, among many other processes, parsing the syntactic structure of a given sentence, resolving ambiguities, inferring the most likely referent of pronouns, etc. To distinguish between theories, reading studies rely on a linking hypothesis that connects the reading time at a word to its processing difficulty rather intuitively: reading times increase monotonically with processing difficulty. While widespread, this linking hypothesis is left implicit or, at best, stated informally without any explicit model.

A major problem with this intuitive linking hypothesis between reading times and processing difficulty is that it cannot explain spillover effects. Spillover effects refer to the observation that characteristics of word n might affect the reading times of words n+1, n+2, etc. This raises concerns about the adequacy of the simple link between processing and reading times. This is important because an inadequate link might make it challenging to discover the mechanisms underlying sentence processing: it is not unlikely that different readers or different stimuli will lead to different spillover effects washing out the effect of an experimental manipulation. In addition, another concern with this linking hypothesis is that it implies a strict serial processing in reading that has been recently questioned (e.g., Engbert et al, 2005, and Snell & Grainger, 2019).

In this talk, I will discuss the limitations of current computational models of reading to account for psycholinguistic phenomena and self-paced reading data. I will describe the characteristics of a novel (self-paced) reading computational model which makes the linking hypothesis connecting latencies and processing explicit. In this model, reading times are a consequence of an autonomous timer (as in SWIFT, Engbert et al, 2005, and ICAT, Trukenbrod & Engbert, 2014) which is modulated by the continuous flow of information (e.g., Townsend & Wenger, 2021) that cascades across different levels of representation, from the visual input to the sentence-level representation.

Click here to see the colloquium program.