Cueing syntax during reading: Eye-movement evidence for benefits of linguistically-driven text formatting

[Authors]

Background. Syntactic structures are fundamental to the construction of meaning, however syntactic composition can be more difficult during reading compared with listening, since many cues like gestures or prosodic phrasing may be lacking. In addition, there is a high occurrence of poor reading comprehension among individuals with poor syntactic competence, such as language learners or those with developmental disabilities. Previous research demonstrates comprehension benefits from alternative text formats, such as presenting sentences in multi-word “chunks”; however, these methods were often implemented in an ad hoc fashion with inconsistent chunking strategies. Linguistically-driven text formatting (LDTF or “Cascade”) is an NLP-driven algorithm that renders linguistic structures transparent by transforming sentences into constituents, with hierarchical relationships indicated via indentations (Fig.1). By aligning subject and verb, indenting dependent information past its head, and indenting elements of a conjunction past their conjoining term, LDTF produces a visual map to sentence structure. Previous research in elementary schools (Authors, submitted), and adult L1 and L2 English readers (Authors, 2024) demonstrates that LDTF improves comprehension during reading.

Current Study. To investigate processing differences between reading traditionally formatted vs. LDTF-formatted English texts, we conducted a 3-session reading and eye-tracking study with 80 adult L1 English readers. Participants were randomly assigned to the LDTF or control group. Session 1 contained a language history questionnaire, spelling recognition task, English reading comprehension task, and read-aloud task to determine baseline English reading abilities, added as controls in inferential models. In Session 2, participants read two 1000-word passages (Fig.4), presented one sentence at a time, each followed by a series of comprehension questions probing basic information from the text (Fig.1). answering Y/N questions after each. These materials were formatted according to experimental group (traditional vs LDTF). In addition, those in the LDTF group read several slides at the beginning of Session 2 describing the indenting principles behind the LDTF format; the control group saw no introductory material. Session 3 was identical to Session 2 except that participants’ eye-movements were recorded.

Results. All models were fit to hierarchical Bayesian models with mildly informative priors and maximal random effects structures by item and participant and were evaluated based on 89% credible intervals. Data from passage reading accuracy in Sessions 2 and 3 were fit to a Bernoulli distribution with fixed effects of Format, Session, and their interaction, revealing higher accuracy for the LDTF format. For eye-tracking measures on passages, beneficial effects were observed for LDTF, including less time spent rereading overall, higher overall skipping rates, and less time spent reading the sentence-final region (Fig.2). These findings suggest that LDTF facilitates the reading process, leading to less rereading and more skipping, perhaps indicative of higher structural certainty, all while improving overall reading comprehension. Results are discussed within the broader frameworks of implicit learning and chunk-and-pass processing, with consideration given towards educational interventions focusing on syntactic knowledge and sentence-level understanding as a key component of the reading comprehension process.

Figure 1. Principles behind LDTF with an example sentence.

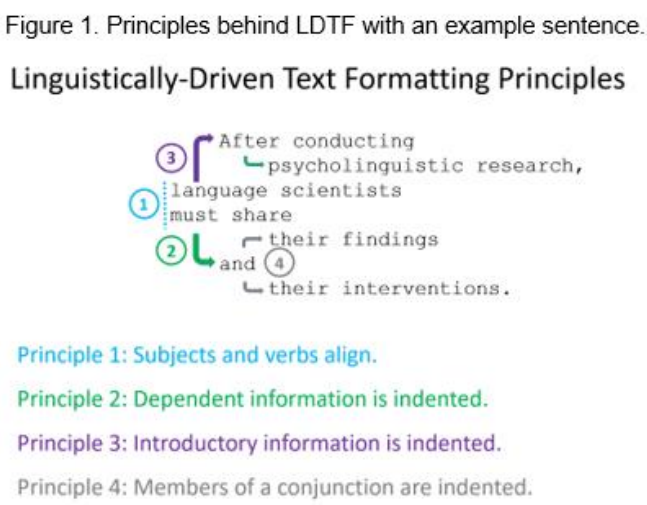

Figure 4. Example of LDTF applied to a reading passage

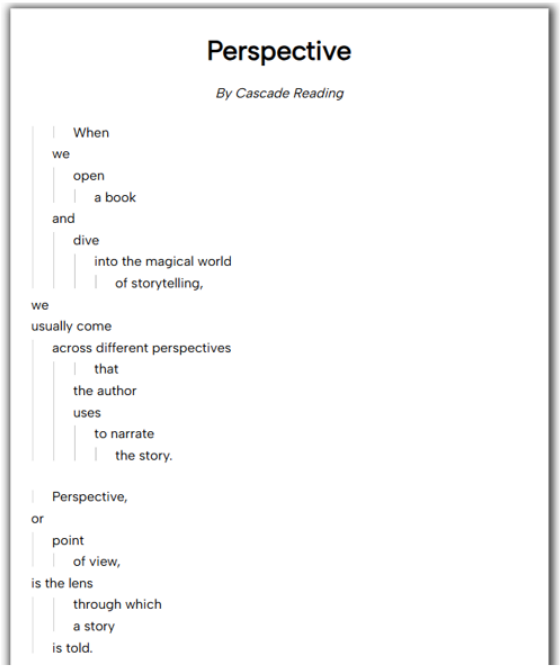

Figure 2. Aggregate reading behavior means by participant across Format groups.

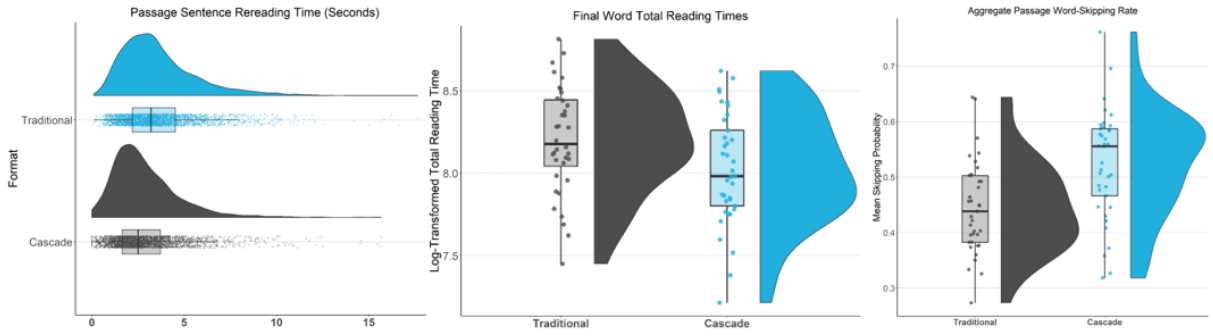

Figure 3. Aggregate mean regression out probability by participant across Format.