

Effects of dependency length on the processing and understanding of texts

Suzanne Kleijn, Henk Pander Maat & Ted Sanders
Utrecht University - S.Kleijn1@uu.nl

Introduction

SYNTACTIC DEPENDENCY LENGTH (SDL) is a measure of the number of words between a syntactic head and its dependent (e.g., verb - subject).

- Longer SDLs require more mental resources from the reader. This is reflected in longer reading times ('Locality effect'; Gibson, 1998; 2000; Demberg & Keller, 2008; Bartek et al. 2011).
- However, interposing elements may guide predictions for upcoming materials and may in fact facilitate processing ('Anti-locality effect'; Konieczny, 2000; Vasishth & Lewis, 2006).
- In addition: text comprehension is not always affected (Gibson, 1998; Renkema, 1991).

For languages that allow some variation in word order (like Dutch) this raises the question whether we should strive to minimize SDLs in order to aid our readers or not.

RQ: How does SDL influence the on-line processing and comprehension of Dutch texts?

Method

- 47 Dutch 9th grade pre-vocational students (lowest level of Dutch educational system) read 4 texts.
- Their eye movements were recorded while reading (Eye-link 1000 eye-tracker; @500Hz).
- Each text was followed by 8 multiple choice questions to measure overall text comprehension.

Materials

- 4 Dutch real-life public information texts (300 - 400 words).
- SDL was increased or decreased by changing the word order in 1/3rd of the sentences resulting in a 'short SDL' and a 'long SDL' version of each text (see Examples below).
- Different types of SDL were manipulated, depending on the possibilities offered by the specific sentence (e.g., verb - subject; verb - object; finite verb - non-finite verb).
- Potential confounding factors were kept constant between text versions (e.g. sentence length, number of sentences, meaning/included information, coherence, discourse focus).
- Sentences were not presented in isolation but in their natural context.

Measures

Name	Description
First pass gaze duration (FPG)	Summed duration of all fixations and intermittent saccades within a sentence in first pass before the eyes leave the sentence (either regressively or progressively)
First pass total gaze duration (FPTG)	Summed duration of all fixations and intermittent saccades within a sentence in first pass before the eyes leave the sentence progressively
First pass regression path duration (RP)	Summed duration of all fixations and intermittent saccades within a sentence in first pass plus regressions to previous sentences before the eyes leave the sentence progressively
Total fixation duration (TFD)	Summed duration of all fixations within a sentence (including second, third... n th pass)
Fixation count FPG/FPTG/RP/TFD	Number of fixations made during First pass gaze/First pass total gaze/First pass regression path/Total fixation
Comprehension score	Score given to each question (1 = correct; 0 = incorrect)

Examples

Short SDL version

(1) a. De aangehoudene dient onverwijld overgedragen te worden aan een opsporingsambtenaar (politie).
The detainee needs immediately handed over to be to a criminal investigator (police).
'The detainee needs to be handed over to a criminal investigator (police) immediately.'

Short SDL version

(2) a. Als je diabetes hebt, zit er te veel suiker in je bloed.
If you diabetes have, is there too much sugar in your blood.
'If you have diabetes, your blood contains too much sugar.'

Long SDL version

b. De aangehoudene dient onverwijld aan een opsporingsambtenaar (politie) overgedragen te worden.
The detainee needs immediately to a criminal investigator (police) handed over to be.
'The detainee needs to be handed over to a criminal investigator (police) immediately.'

Long SDL version

b. Als je diabetes hebt, zit er in je bloed te veel suiker.
If you diabetes have, is there in your blood too much sugar.
'If you have diabetes, your blood contains too much sugar.'

Results

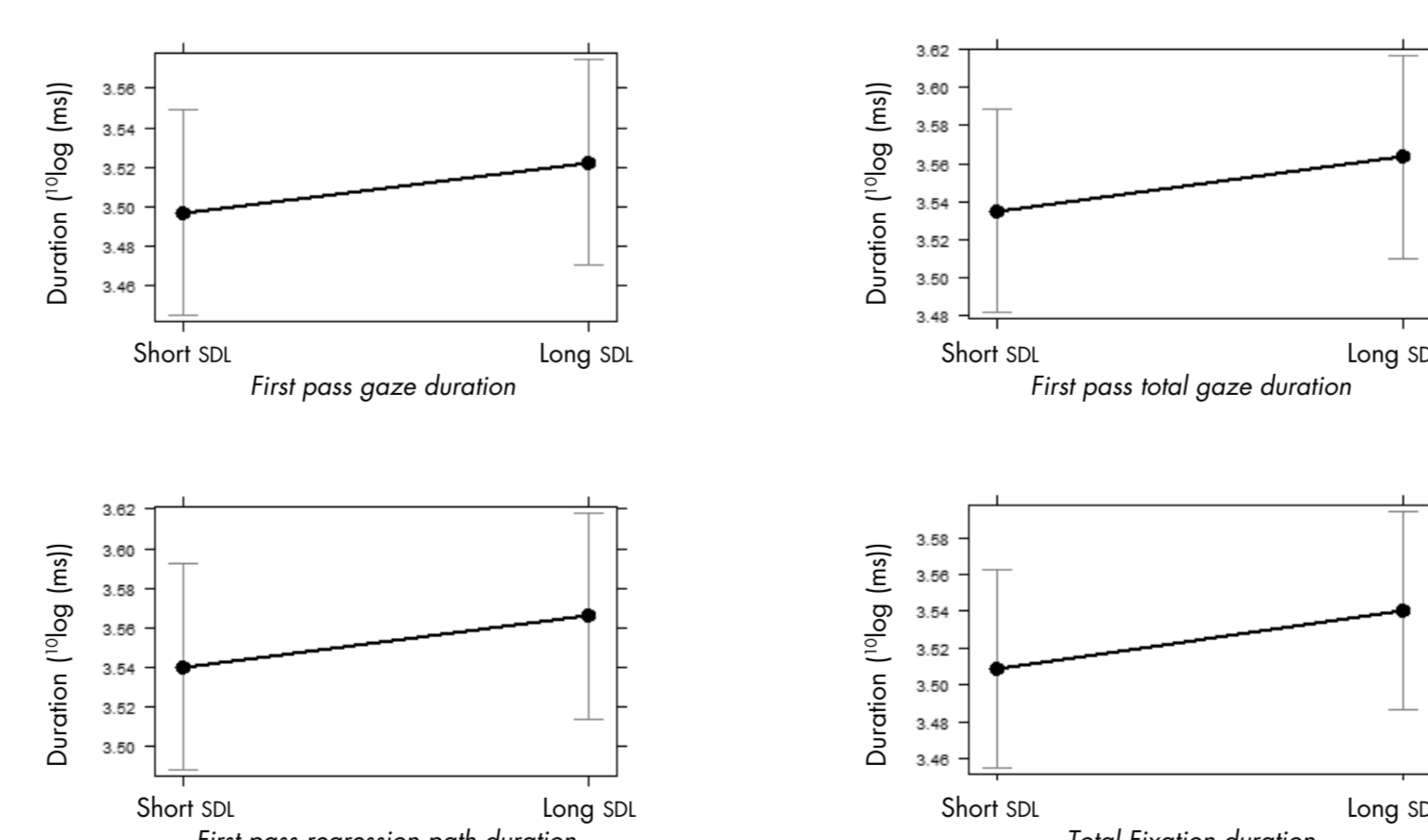
Sentence reading times

- Linear mixed effect modeling
 - Random effects: Subject, Sentence & Text
 - Fixed effects: Text version, mean length of words in the sentence (in letters), sentence length (in words)
 - N.s. effects: standardized reading ability scores, trial
- Sentence reading times were higher for manipulated sentences in the 'long SDL' version compared to the 'short SDL' version (FPG, FPTG, RP and TFD)
- Number of fixations was higher in the 'long SDL' version sentences compared to the 'short SDL' version (FPG, FPTG, RP and TFD)

Comprehension scores

- Generalized mixed effect modeling
- No effect of SDL on comprehension scores

Fixed effects of SDL



Conclusion & Discussion

- The results are in line with the locality effect: Increasing syntactic dependency lengths increases sentence processing times.
- Word level analysis must reveal the specific time course of the effect.
 - Facilitating effect of intervening materials may still happen!
- Even pre-vocational students are able to overcome the additional processing demand of longer SDLs. Processing does not break down and comprehension does not seem compromised.

References

- Bartek, B., Lewis, R. L., Vasishth, S. & Smith, M. R. (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1178-1198.
- Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109(2), 193-210.
- Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68(1), 1-76.
- Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz & W. O'Neil (Eds.), *Image, language, brain* (pp. 96-126). Cambridge: MIT Press.
- Hall, T. (2006). *Eye* [Online image]. Retrieved from: <https://flic.kr/p/mrF5E>. Used under CC BY-NC-SA 2.0 / Desaturated from original.
- Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic research* 29(6), 627-645.
- Renkema, J. (1991). Text quality, three approaches to experimental research. *Revue Belge de Philologie et d'Histoire*, 69(3), 618-628.
- Vasishth, S., & Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 82(4), 767-794.