

Dependency Parsing with your Eyes: Dependency Structure Predicts Eye Regressions During Reading

Alessandro Lopopolo

Center for Language Studies
Radboud University
Nijmegen, the Netherlands.
a.lopopolo@let.ru.nl

Stefan L. Frank

Center for Language Studies
Radboud University
Nijmegen, the Netherlands.
s.frank@let.ru.nl

Antal van den Bosch

Meertens Institute - KNAW
Amsterdam, the Netherlands.
Center for Language Studies
Radboud University
Nijmegen, the Netherlands.
a.vandenbosch@let.ru.nl

Roel M. Willems

Center for Language Studies & Donders Institute
Radboud University
Nijmegen, the Netherlands.
MPI for Psycholinguistics
Nijmegen, The Netherlands.
r.willems@donders.ru.nl

Abstract

Backward saccades during reading have been hypothesized to be involved in structural re-analysis, or to be related to the level of text difficulty. We test the hypothesis that backward saccades are involved in online syntactic analysis. If this is the case we expect that saccades will coincide, at least partially, with the edges of the relations computed by a dependency parser. In order to test this, we analyzed a large eye-tracking dataset collected while 102 participants read three short narrative texts. Our results show a relation between backward saccades and the syntactic structure of sentences.

1 Introduction

Written language consists of a sequence of graphic signs. While most eye movements during reading indeed follow this sequential order, they are also occasionally interleaved with jumps back to words in preceding portions of the text. We refer to these backward saccades as regressions throughout this paper. There are at least two competing hypotheses concerning the nature and function of this phenomenon.

The first main line of hypotheses on the role of regressions emphasizes their active role in computing linguistic representations (Kennedy, 1992), the second stresses their function as a reanalysis tool in the event of detected parsing errors (Rayner and Sereno, 1994). In this paper we start from the former; in particular, we aim to investigate the relation between regressions and the structure of sentences as computed by a dependency parser.

We take dependency structures as a valid approximation of syntactic properties of the sentences, and we investigate whether these are reflected in eye movement regressions during naturalistic text reading.

We consider regressions from each word in the text, and relate those to dependency relations that link pairs of words in the sentence. In this way we can represent syntactic properties of the sentences as shallow structural information at the word level, by focusing on the number and direction of the syntactic relations that each word in a sentence is engaged in.

The aim of this paper is two-fold: on the one hand we want to investigate whether regressions might play a role in online sentence parsing; on the other hand – as an implication of the previous goal – we are interested in finding traces of syntactic parsing during reading. We report the results of a mixed-effect regression analysis showing a relation between the pattern of eye regressions and the syntactic structure of sentences.

2 Theoretical Background

2.1 The Role of Regressions in Text Comprehension

Regressions (backward saccades) are relatively rare, occurring usually only with 15 to 25% of the words (Rayner & Pollatsek, 1989). They do not seem to be random, however. Regressions typically aim at specific word locations, moving fixation from the current word back to one of the previously encountered words (Vitu, 2005). Nonethe-

less their function in language comprehension is still debated. Here we will review two proposed explanations: the first links regressions to the difficulty of text processing; the second instead sees them as tools for language processing, not necessarily linked to processing difficulties or errors. According to the first proposal regressions only start to play a role in reading once difficulties are encountered; according to the second proposal they are part and parcel of regular reading.

2.1.1 Regression as a Response to Comprehension Difficulty

The first hypothesis interprets regressions as part of the reanalysis of textual input due to encountered comprehension problems. In a milestone study, [Altmann et al. \(1992\)](#) introduced the notion of regression-contingent analysis, based on the assumption that regressive eye movements are a necessary consequence of subjects being garden-pathed. A garden-path effect occurs when readers incrementally construct an incorrect interpretation of a sentence as a consequence of its locally ambiguous syntactic structure. This does not necessarily mean that the presence of a difficult structure, leading for instance to the reader being garden-pathed, triggers a regression. Rayner and colleagues reported data showing that strong garden path effects can occur sometimes without triggering any regressions ([Rayner and Sereno, 1994](#); [Castelano and Rayner, 2008](#)). Nonetheless, other studies have given support to the idea that regressions are linked to textual ambiguity and contextual difficulties. Readers make more regressions when the text is complex ([Rayner and Pollatsek, 1995](#)), when the topic changes ([Hyn, 1995](#)), when the text contains grammatical errors or ambiguities ([Reichle et al., 2013](#)), or when they encounter information that disambiguates the preceding text ([Blanchard and IranNejad, 1987](#); [Frazier and Rayner, 1982](#)). The general hypothesis holds that the probability of regressions and their span might depend on the difficulty of the text. Therefore these regressions might allow the reader to reread information that has been missed, forgotten, or wrongly interpreted ([Rayner, 1998](#)).

2.1.2 Regression as a Tool for Comprehension

The alternative explanation focuses instead on the role of eye movements as a tool in language processing, used independently from the structural

difficulty of the input. The idea is that regressions help the reader reactivate cognitive information that is associated with the regressed-to location. [Kennedy \(1992\)](#) refers to this as the Spatial Code Hypothesis. The hypothesis is that readers use the position of words on the page as a support to their working memory by reactivating previously read words associated with information relevant for the processing of the word from which the regression originated ([O'Regan, 1992](#); [Spivey et al., 2004](#)). This hypothesis has been criticized by [Booth and Weger \(2013\)](#). They presented three experiments showing that reader comprehension is not hindered when reading conditions inhibit or discourage visual access to already read material. In their Experiment 1, readers knew that candidate targets for regression were no longer available for rereading. Experiment 2 discouraged regressions by forcing readers to follow a visual placeholder on the stimulus while it was also presented in auditory form. Finally, in Experiment 3, candidate targets for regression were manipulated after reading. In all these three experiments, readers showed no hindered comprehension of the presented stimulus sentences. As an entailment of these results, Booth and Weger suggested that readers do not use regressions to cue their memory for previously read words.

Our hypothesis is that readers might make use of regressions to reactivate previously read information in the context of naturalistic language comprehension, in order to help compute linguistic information.

We want to examine whether there is an alignment between patterns of regressions and word-to-word syntactic relations as described by the dependency structure of the stimulus. We hypothesize that regressions play a role in syntactic parsing that may go beyond the reanalysis of ambiguous material. We do not deny their role in reanalysis and repair, but we rather stand with the hypothesis that they allow re-reading and cueing of previous words, as an aid to memory, when this is required for a successful construction of a syntactic representation of the text.

In order to test this hypothesis we rely on an eye-tracker dataset that was collected during normal text reading of unmodified literary narratives. We assess whether there is a relation between the number of eye regressions from the words and the number of syntactic relations that those words en-

tain with their preceding text. These syntactic relations are derived from the dependency structures (described in Section 4) of the sentences composing the stimuli of the eye-tracker dataset.

2.2 Regressions and Sequence Processing

The hypothesized relation between dependency structure and eye movement taps into a broader debate on whether language processing relies mainly on the sequential structure of the input or whether it involves the computation of non-sequential syntactic parses (Jackendoff and Wittenberg, 2014). Undeniably, the linguistic stimulus is presented as a string of symbols, nonetheless regressions seem to counter the notion that it is processed strictly in a sequential order. If these eye movements are involved only in re-analysis, then their existence does not necessarily contradict sequential processing accounts. They can be explained as an "emergency recovery" operation that takes place only in cases of processing difficulties. On the other hand, if we find evidence of a relation between saccades and syntactic dependency structures independent from processing difficulty, then we might conclude that saccades offer behavioral evidence that processing involves the computation of non-sequential structures.

This question is related to the line of research in psycholinguistics and neuroscience investigating the computation of syntactic structures in the mind/brain during language processing. In this context, sequential structures are usually contrasted with hierarchical ones, where input items are grouped into larger units, which in turn are (possibly recursively) grouped in even larger units. These larger units are commonly referred to as syntactic constituents or phrases and have a central position in theoretical linguistics (Chomsky, 1965; Jackendoff, 2002, 2007). An increasing amount of evidence against a strictly hierarchical processing of language has emerged over the past decades. Psycholinguistic studies have supplied evidence suggesting that the mere sequential properties of the stimulus are sufficient to explain aspects of human behavior during reading and listening. Frank et al. (2012) provide a review of evidence from cognitive neuroscience, psycholinguistics and computational modeling studies supporting the hypothesis that hierarchical structure may not play a central role in language processing and acquisition, and that sequential structure

instead has a significant explanatory power. They argue that hierarchical structure is rarely needed to explain behavioral and neural correlates of language processing in vivo. In contrast with these findings, recent neuroimaging studies have delineated a slightly more complex landscape in which both hierarchical and sequential processing may be carried out simultaneously by the human brain during language processing (Brennan et al., 2016; Nelson et al., 2017).

Dependency parses are different from constituency parses as they lack the non-terminal nodes characteristic of constituency parses. Nonetheless they still constitute a non-sequential type of structure. Demonstrating a relation between eye movement and such structure will provide evidence for the non-sequentiality of language processing, at least in the context of text reading.

3 Related Work

The present work studies the relation between eye movements during reading and the dependency structure as produced by a dependency parser (see Section 5.2 for more details). Several other studies tested language processing hypotheses by using computational models as predictors of eye movements during sentence reading.

Boston et al. (2008) demonstrates the importance of including parsing costs implemented as surprisal as a predictor of comprehension difficulty in models of reading. They showed that surprisal of grammatical structures has an effect on fixation durations and regression probabilities.

Demberg and Keller (2008) compared linguistic integration cost computed as a function of dependency relations distances and word surprisal as predictors of gaze duration. They showed that distance is not a significant predictor of reading times except for nouns. On the other hand, they demonstrate that surprisal can predict reading times for arbitrary words in the corpus, concluding that the two predictors may capture distinct aspects of naturalistic language processing.

In the context of Natural Language Processing, Klerke et al. (2015) used eye-tracker data as a metric for the quality of automatic text simplification and compression, which are operations used in machine translation and automatic summarization. Their proposal is grounded in the hypothesis that eye movements are related to perceived text diffi-

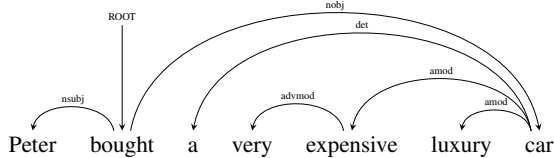


Figure 1: Dependency Parse of Sentence 1

culty (Rayner and Pollatsek, 1995), one of the two hypothesis we have introduced in Section 2 above.

4 Dependency Structure

We chose dependency grammar as the formalism of non-sequential syntactic structure. Dependency grammar describes a sentence as a set of relations between words (heads) and their dependents. These relations are called dependencies and correspond to grammatical functions and – together with the words they link – are the only descriptive elements composing the structure, which has the form and properties of a directed graph (Tesnière et al., 2015; Mel’čuk, 1988; Nivre and Kübler, 2009).

- (1) Peter bought a very expensive luxury car.

Take for instance Sentence 1 above. Figure 1 contains the dependency graph representing the dependency structure of the sentence in terms of typified head-dependent relations. The main verb (*bought*) acts as head for *Peter* and *car*, with which it is in **subject** and **object** relations, respectively. A dependent of one dependency relation can in turn be the head of another one. For instance *car* is head of *luxury* and of *expensive* with which it is linked by **modifier** relations, and also head of the article *a* via a **determiner** relation.

This structure lacks phrasal non-terminal constituents. In addition, it is not strictly sequential, or put differently, it is not isomorphic to the sequence of items that makes up the stimulus. This is based on the fact that the dependency relations can hold between words that are non-consecutive or possibly even far apart in the sentence. There is the assumption that during reading, these links are created once a suitable candidate for the second term of the dependency is introduced. Therefore, online dependency parsing proceeds by introducing one word at a time, and by looking back at the prefix in order to assess whether this novel input is a suitable candidate for a dependency link with a preceding word that has not yet been matched.

5 Materials and Methods

5.1 Eye-tracker data

The eye tracker data used in this study was originally collected for a study on mental stimulation during literary reading by Mak and Willems (2018) at Radboud University, Nijmegen, the Netherlands. For more details on data acquisition and preprocessing we refer to the original publication.

5.1.1 Participants and Stimuli

Data was collected from 102 participants (82 females, mean age 23.27, range 18–40), all of whom were native speakers of Dutch, with normal or corrected-to-normal vision. All participants gave written informed consent in accordance with the Declaration of Helsinki.

Stimuli consisted of three published short stories in Dutch. Stories 1 and 2 were written by contemporary Dutch writers, and Story 3 was translated from American English to Dutch. Their lengths were 2143, 2659, and 2988 words respectively, and they required around 10-15 minutes each to be read.

5.1.2 Data Acquisition and Pre-processing

For eye-movement data collection, a monocular desktop-mounted EyeLink1000+ eye-tracking system was used (500 Hz sampling rate). Head movements were minimized using a head stabilizer, ensuring that all participants were seated at 108 cm from the screen.

The stimuli were presented using SR Research Experiment Builder software (SR Research, Ottawa, Canada). The stories were divided into 30 sections each. The stories were presented in counterbalanced order. After data collection, participants were presented with a comprehension questionnaire.

All fixations were checked to make sure that they did not drift off and enter a different interest area. If correction of the drifts was not possible, individual sections were excluded. Data for at least one section was removed for 40 participants. For four participants, the number of excluded sections exceeded six, resulting in the exclusion of one story for these participants.

Eight participants answered more than one comprehension question incorrectly for one of the three stories (four times for Story 2 and four times

for Story 3), resulting in the exclusion of the data for one story reading for eight participants.

The dataset contains a total of 582,807 words across all participants and narratives.

5.1.3 Eye Tracker Measures

For the present study we focus on the number of eye regressions. A regression consists of a fast eye movement from a word back to a previous word.

5.2 Dependency Parsing

The text of the three stories presented to the participants were fed to the ALPINO toolbox for Dutch natural language processing (Noord, 2006) to generate a dependency parse for each of their sentences. The parser creates a structure composed of dependency triples consisting of a head word, the type of dependency relation and its dependent word. A parse is produced for each sentence independently, therefore no relation can be assigned between words belonging to different sentences. The output of the parser was manually checked in order to prevent tokenization and sentence segmentation errors.

5.2.1 Number of Dependency Relations

As described in Section 4, every word in a sentence entertains at least one relation with another word in the same sentence. Every non-final and non-initial word can have relations with a variable number of other words on its right and its left. Because we are interested in eye regressions, we decided to focus our attention only on relations between a word w and its preceding context. Therefore only relations with a head and possible dependents on the words preceding w are counted.

From the dependency structure of a sentence, we derived the following count measures:

- ***N_head*** indicates the presence of a syntactic relation between w_i and a word in $w_{1:i-1}$ that is head of w_i ;
- ***N_dependents*** counts the number of syntactic relations between w_i and words in $w_{1:i-1}$ that are dependents of w_i .

Measure *N_head* is a binary variable indicating whether word w has a head in its left context $w_{1:t-1}$. This is because every word has one, and only one, head.

For example, the word *expensive* in Sentence 1 has one head relation with a word on its right

	Peter	bought	a	very	expensive	luxury	car
<i>N_head</i>	0	0	0	0	0	0	1
<i>N_dependents</i>	0	1	0	0	1	0	3

Table 1: Number of dependency relations per word w that correspond to words in w 's own preceding context.

(*car*), no dependents on its right, and one on its left (*very*). On the other hand, the word *car*, being sentence-final, does not have any links on its right, but it has 1 head (*bought*) and 3 dependents (*a*, *expensive*, *luxury*) on its left. Table 1 contains the count measures (*N_head* and *N_dependents*) for Sentence 1.

5.3 Descriptors not Related to Dependencies

We are interested in the effect of syntactic structure, implemented as dependency relations, on the pattern of regressions. For this reason it is necessary to control for other possible quantifiable factors affecting these eye movements. We chose to use log-transformed lexical frequency and surprisal.

Base-2 log-transformed lexical frequency per word was computed using the Subtlex NL corpus (Keuleers et al., 2010). Surprisal was computed from a second-order Markov model, also known as trigram model, trained on a random selection of 10 million sentences (comprising 197 million word tokens; 2.1 million types) from the Dutch section of Corpora from the Web (NLCOW2012; Schäfer and Bildhauer, 2012). Surprisal of word w_t is the negative logarithm of the conditional probability of encountering w_t after having read sequence w_{t-2}, w_{t-1} , or: $-\log P(w_t | w_{t-2}, w_{t-1})$. The computation was performed by the SRILM toolbox (Stolcke, 2002).

Frequency and surprisal are computed in order to control for processing difficulties. Intuitively, infrequent words and words with high surprisal are more difficult to retrieve (and possibly to integrate) with their preceding context. Controlling for processing difficulty is motivated by the alternative hypothesis regarding the role of regressions as depending on the level of complexity posed by a linguistic input.

In addition to frequency and surprisal, we also use word position in the sentence. Intuitively, the probability of regressing from a word to its previous context increases linearly with the position of the word in a sequence. By controlling for it, we ensure that the eye movements are not due simply

to the opportunity given by the larger target pool to regress to.

5.4 Analyses

We fitted two logistic mixed-effect models predicting eye regressions. The first model (**null**, Eq. (1) below) contains as predictors only the position of the words in their sentences (*word_order*), and probabilistic information consisting of the above-mentioned log-transformed frequency (*freq*) and surprisal (*surp*). The second model (**full**, Eq. (2) below) contains as predictors of interest also the number of left-hand side dependency relations (i.e. *N_head* and *N_dependents*) of each word. In addition, we included by-participant and by-word random intercepts, as well as by-participant random slopes for *word_order* in the **null** model and for *word_order*, *N_head* and *N_dependents* in the **full** models.

We expect the model’s fit to improve significantly after inclusion of the measures derived from the dependency parse as regressors. The increase in model fit is quantified by the χ^2 -statistic of a likelihood-ratio test for significance between the **null** and **full** models and is taken as the measure of the fit of *N_head* and *N_dependents* measures at each word to the probability of a regression being generated at each word.

$$\begin{aligned} \text{null} : \text{eye_regressions} \\ \sim \text{word_order} + \text{surp} + \text{freq} + (1|\text{word}) \\ + (1 + \text{word_order}|\text{participant}) \end{aligned} \quad (1)$$

$$\begin{aligned} \text{full} : \text{eye_regressions} \\ \sim \text{word_order} + \text{surp} + \text{freq} \\ + N_head + N_dependents \\ + (1|\text{word}) + (1 + \text{word_order} + N_head \\ + N_dependents|\text{participant}) \end{aligned} \quad (2)$$

The models are fit by maximum likelihood (Laplace Approximation) and with a binomial distribution.

6 Results

6.1 Regression Model Analysis

In the results below we first describe the fit of each of the two models (**null** and **full**) separately, then we report the results of the model comparison analysis using the χ^2 -statistic.

Table 2 presents the fitted **null** model. Table 3 shows the fitted **full** model. The head and dependent regressors have significant effects on the number of regressions (*eye_regressions*) - *N_head*: $\beta = 0.242, p < .0001$; *N_dependents*: $\beta = 0.046, p < .0005$.

In addition, both word frequency (*freq*) and surprisal (*surp*) have a significant negative effect. The negative effect of frequency might be due to less frequent words being more difficult to retrieve from memory, therefore triggering a regression to gather more contextual information to help word processing. The negative effect of surprisal indicates that the larger the surprisal of a word – therefore more difficult its integration into the context – the less probable the reader is to regress to the word’s previous context. Mak and Willems (2018) reported a positive effect of surprisal on the number of incoming saccades, that is, eye movements into a word back from subsequent parts of the text.

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.616	0.049	-32.9	< .0001
word_order	0.035	0.007	5.2	< .0001
surp	-0.140	0.013	-11.1	< .0001
freq	-0.165	0.028	-6.0	< .0001

Table 2: Fixed effects for the **null** model

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.798	0.049	-36.6	< .0001
word_order	0.019	0.006	3.0	< .003
surp	-0.102	0.012	-8.0	< .0001
freq	-0.125	0.027	-4.6	< .0001
N_head	0.242	0.016	14.8	< .0001
N_dependents	0.046	0.013	3.6	< .0005

Table 3: Fixed effects for the **full** model

In order to test whether the introduction of head and dependent measures improves the fit of the logistic mixed effect model to outgoing saccades, we computed the χ^2 -statistic of a likelihood-ratio test for the difference between the **null** and **full** models above. The χ^2 is taken as the measure of the fit of the dependency measures to the probability of a regression being initiated at each word. Table 4 reports the results of the test, showing the difference in model fit to be significant ($\chi^2 = 738.87, p < .0001$).

6.2 Analysis of Regression Counts

The results of the regression model comparison indicate that regressions are partially driven by

model	df	AIC	BIC	deviance	χ^2	χ^2 df	Pr
null	8	401023	401111	401007			
full	17	400302	400489	400268	738.87	9	< .0001

Table 4: Results of log-likelihood comparison between **null** and **full** model.

the presence of left-hand side dependency relations. In order to corroborate these observations, we counted the number of times regressions generated from each word do actually land on preceding words that are heads or dependents of that word. As reported in Table 5, it turns out that of the 110,336 regressions, about 40% do actually land on a head or dependent of the words they originate from. These are referred to as matches. The analyses were limited only to regressions landing within sentence boundaries. In the table, “misses” refers to the regressions that land on targets that are neither head nor dependent of the the word they originated from.

tot nr of regressions:	110336
tot nr of matches:	46378
tot nr of misses:	63958

Table 5: Total numbers of regressions, matches (i.e. regressions that land on heads or dependents), and misses (i.e. regressions do not land on heads or dependents of the word they originated from).

A χ^2 -test of independence was performed to assess the relation between having a dependency relation with a word and generating a regression to that word. The test was computed independently for 10 separate left-hand side distances $d = [-10 : -1]$. In other words, for $d = -1$, we want to assess whether there is a relation between having a dependency relation with the preceding word and looking back at that word; for $d = -2$, we want to assess whether there is a relation between having a dependency relation with the preceding word at position -2 and looking back at that word, and so on for the other considered distances.

Table 6 contains the per-distance results of the χ^2 analyses. An association between presence of a dependency relation at position d and the generation of a regression to that position is significant for distances -1 ($\chi^2 = 132.52, p < 0.001$), -2 ($\chi^2 = 678.14, p < 0.001$), -3 ($\chi^2 = 8.05, p < 0.005$), and -4 ($\chi^2 = 13.68, p < 0.001$). For all other tested distances (between -5 and -10) the association was not significant (see Figure 2). For

$d = \{-1, -2, -3, -4\}$:

- The fraction of words w_i in a dependency relation with w_{i-d} that originate a regression of length $-d$ is significantly higher than the fraction of w_i not in a dependency relation with w_{i-d} originating a regression of length $-d$;
- The fraction of words w_i with a regression of length $-d$ that are in a dependency relation with w_{i-d} is significantly higher than the fraction of w_i without a regression of length $-d$ that are in a dependency relation with w_{i-d} .

dist	+dp+reg	-dp+reg	+dp-reg	-dp-reg	χ^2
-1	29245	19520	305931	228111	132.52**
-2	5711	14798	113028	449270	678.14**
-3	937	6239	68785	506846	8.05*
-4	309	2641	49324	530533	13.68**
-5	76	1370	33153	548208	0.55
-6	35	850	29823	552099	2.69
-7	17	530	23728	558532	1.42
-8	13	355	17687	564752	0.29
-9	5	261	16774	565767	1.08
-10	3	263	13785	568756	2.18

Table 6: χ^2 analyses for distances $-10 : -1$. **+dp+reg** indicates the number of words in the corpus having a relation and a regression at $-d$; **+dp-reg** the number of words having a relation but not a regression at $-d$; **-dp+reg** number of words not having a relation but having a regression at $-d$; **-dp-reg** not having nor relation nor regression at $-d$ (** = $p < 0.001$, * = $p < 0.01$).

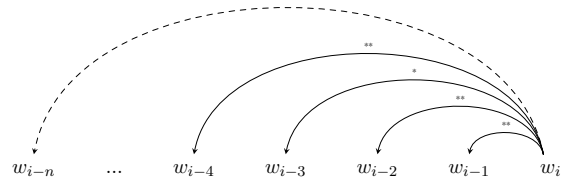


Figure 2: The effect of dependency relations on regressions from w_i is significant only for the preceding four words - further away saccades might not be influenced by a possible relation with w_i (** = $p < 0.001$, * = $p < 0.01$).

This seems to indicate that the effect of the dependency structure of a sentence on the pattern of outgoing eye-movements from a word is present only for short-distance relations (between a word and its four preceding words).

It is important to keep in mind however that the number of dependency relations found by the

parser is much higher than the actual number of matches. This is simply because the parser does assign at least a head to each and every word in the text (even words in isolation are assigned a root head), whereas a regression is a relatively rare event (under normal conditions, using naturalistic language). The present work aims at demonstrating that regressions are related (also) to the structure of the dependency graph. The results we have produced so far point in that direction. In other words, it is possible to affirm that if a regression takes place, it might be triggered by the presence of a dependency relation between the word it is generated from and the word it lands on.

7 Conclusion

In this paper we have presented an analysis investigating whether eye movements of readers may be influenced directly by the syntactic structure of the sentence. We tested this using shallow non-hierarchical structures computed by a dependency parser. The hypothesis was that the path of regressions from a word to an earlier word coincide, at least partially, with the edges of dependency relations between these words. We used a substantially large eye-tracking dataset collected while 102 participants were engaged in reading three short narrative texts.

The results of a mixed-effect regression analysis indicate that there is a significant effect of the number of left-hand side dependency relations on the number of backward saccades. This effect is well above chance even after correcting for word position in the sentence and word frequency and surprisal – measures held to explain a large part of natural language processing behavioral and neural correlates. These results are corroborated by the observation that about 40% of backward saccades do indeed land on target words engaged in dependency relations according to the syntactic structure of the sentences composing our stimuli.

The length of the regressions seems to be relatively short: the vast majority being shorter than three words, with a predominance of regressions one position backwards. The results of a series of χ^2 tests for independence shows that there is a significant association between presence of a dependency link and backward saccading between two words holds only for pairs which are not further apart than four positions. This might indicate that the eye regressions are involved

predominantly in dependency parsing at the local level, rather than at long distance.

Altogether these results converge on the idea that eye movements reflect, among other things, the shallow syntactic structure of language. Moreover these results seems also to corroborate the idea that humans do engage in online syntactic analysis of the input – at least in the form of local dependency parsing.

Acknowledgments

The work presented here was funded by the Netherlands Organisation for Scientific Research (NWO) Gravitation Grant 024.001.006 to the Language in Interaction Consortium. The authors thank Marloes Mak for providing the eye-tracker data and help in the analyses.

References

- Gerry T.M Altmann, Alan Garnham, and Yvette Dennis. 1992. [Avoiding the garden path: Eye movements in context](#). *Journal of Memory and Language*, 31(5):685 – 712.
- Harry E. Blanchard and Asghar IranNejad. 1987. [Comprehension processes and eye movement patterns in the reading of surprising stories](#). *Discourse Processes*, 10(1):127–138.
- Robert W. Booth and Ulrich W. Weger. 2013. [The function of regressions in reading: Backward eye movements allow rereading](#). *Memory & Cognition*, 41(1):82–97.
- Marisa Ferrara Boston, John Hale, Reinhold Kliegl, Umesh Patil, and Shravan Vasishth. 2008. [Parsing costs as predictors of reading difficulty: An evaluation using the potsdam sentence corpus](#). *Journal of Eye Movement Research*, 2(1):1–12.
- Jonathan R. Brennan, Edward P. Stabler, Sarah E. Van Wagenen, Wen-Ming Luh, and John T. Hale. 2016. [Abstract linguistic structure correlates with temporal activity during naturalistic comprehension](#). *Brain and Language*, 157–158:81–94.
- Monica Castelhano and Keith Rayner. 2008. Eye movements during reading, visual search, and scene perception: An overview. In *Cognitive and Cultural Influences on Eye Movements*, pages 175 – 195. Tianjin People’s Publishing House.
- Noam Chomsky. 1965. *Aspects of the Theory of Syntax*. The MIT Press, Cambridge.
- Vera Demberg and Frank Keller. 2008. [Data from eye-tracking corpora as evidence for theories of syntactic processing complexity](#). *Cognition*, 109(2):193 – 210.

- Stefan L. Frank, Rens Bod, and Morten H. Christiansen. 2012. [How hierarchical is language use?](#) *Proceedings of the Royal Society B: Biological Sciences*, 279:4522–4531.
- Lyn Frazier and Keith Rayner. 1982. [Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences.](#) *Cognitive Psychology*, 14(2):178 – 210.
- Jukka Hyn. 1995. [An eye movement analysis of topic-shift effect during repeated reading.](#) *Journal of Experimental Psychology Learning Memory and Cognition*, 21:1365–1373.
- Ray Jackendoff. 2002. *Foundations of language: brain, meaning, grammar, evolution.* Oxford University Press.
- Ray Jackendoff. 2007. [A parallel architecture perspective on language processing.](#) *Brain Research*, 1146:222.
- Ray Jackendoff and Eva Wittenberg. 2014. What you can say without syntax: A hierarchy of grammatical complexity. In *Measuring grammatical complexity*, pages 65 – 82. Oxford: Oxford University Press.
- Alan Kennedy. 1992. The spatial code hypothesis. In K. Rayner, editor, *Eye Movements and Visual Cognition. Springer Series in Neuropsychology.* New York: Springer.
- Emmanuel Keuleers, Marc Brysbaert, and Boris New. 2010. SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, 42(3):643–650.
- Sigrid Klerke, Maria Jung Barrett, Sheila Castilho, and Anders Søgaard. 2015. Reading metrics for estimating task efficiency with SMT output. In *Proceedings of the Sixth Workshop on Cognitive Aspects of Computational Language Learning*, pages 6–13. Association for Computational Linguistics.
- Marloes Mak and Roel M. Willems. 2018. [Mental simulation during literary reading: Individual differences revealed with eye-tracking.](#) *Language, Cognition and Neuroscience*, 34(4):511–535.
- Igor Mel'čuk. 1988. *Dependency Syntax: Theory and Practice.* State University of New York Press.
- Matthew J. Nelson, Imen El Karoui, Kristof Giber, Xiaofang Yang, Laurent Cohen, Hilda Koopman, Sydney S. Cash, Lionel Naccache, John T. Hale, Christophe Pallier, and Stanislas Dehaene. 2017. [Neurophysiological dynamics of phrase-structure building during sentence processing.](#) *Proceedings of the National Academy of Sciences*, 114(18):E3669–E3678.
- Joakim Nivre and Sandra Kübler. 2009. [Dependency parsing.](#) *Synthesis Lectures on Human Language Technologies*, 2.
- Gertjan Van Noord. 2006. At last parsing is now operational. In *In TALN 2006*, pages 20–42.
- John Kevin O'Regan. 1992. Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology*, 46:461–488.
- Keith Rayner. 1998. [Eye movements in reading and information processing: 20 years of research.](#) *Psychological Bulletin*, 124:372–422.
- Keith Rayner and Alexander Pollatsek. 1995. *The Psychology of Reading.* Lawrence Erlbaum Associates, Hillsdale.
- Keith Rayner and Sara C. Sereno. 1994. [Regressive eye movements and sentence parsing: On the use of regression-contingent analyses.](#) *Memory & Cognition*, 22(3):281–285.
- Erik D. Reichle, Simon P. Liversedge, Denis Drieghe, Hazel I. Blythe, Holly S.S.L. Joseph, Sarah J. White, and Keith Rayner. 2013. [Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill.](#) *Developmental Review*, 33(2):110 – 149.
- Roland Schäfer and Felix Bildhauer. 2012. [Building large corpora from the web using a new efficient tool chain.](#) In *LREC*, pages 486–493. European Language Resources Association (ELRA).
- Michael Spivey, Daniel Richardson, and Stanka A. Fitneva. 2004. Thinking outside the brain: Spatial indices to visual and linguistic information. In J. Henderson and F. Ferreira, editors, *Interfacing Language, Vision, and Action*, pages 161–190. CA: Academic Press, San Diego.
- Andreas Stolcke. 2002. SRILM - An Extensible Language Modeling Toolkit. In *Proceedings of the International Conference on Spoken Language Processing*, pages 901–904.
- Lucien Tesnière, Timothy J. Osborne, and Sylvain Kahane. 2015. *Elements of Structural Syntax.* John Benjamins Publishing Company.
- Françoise Vitu. 2005. [Visual extraction processes and regressive saccades in reading.](#) *Cognitive Processes in Eye Guidance*, pages 1–32.