The reading process of dynamic text – A linguistic approach to an eye movement study

Selina Sharmin
University of Tampere

Mari Wiklund
University of Helsinki

Liisa Tiittula
University of Helsinki

Abstract

Using eye-movement analysis, the article examines the reading process of speech-to-text interpretation involving dynamic text emerging letter by letter on the screen. The article focuses on regressions of gaze as well as on their relationship to linguistic factors in order to reveal how the reader's gaze behaviour reflects the reading process of dynamic text. The data come from an experiment where participants read a dynamic text on a computer screen. The results showed that the first and second landing points of regressions were generally (90.8%) content words, even though the proportion of content words in the whole data set was only 57.1%. The test subjects looked for nouns, verbs and adjectives in order to construct the meaning of what they had just read. Nouns were the most likely landing points of regressions. The landing points of regressions reflected the reading process through which the meaning of the text was constructed. In this kind of dynamic text, a typical cause of regressions seems to be incoherence resulting from omissions.

Keywords: reading, dynamic text, speech-to-text interpreting, regressions of gaze, eye movements, gaze behaviour, discourse processing, lexical hierarchy

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

1 Introduction

In this article, we examine the reading process of one type of dynamic text: a text that emerges letter by letter on a screen. This kind of text presentation is used in speech-to-text interpretation, where speech is simultaneously rendered into written format. Intralingual speech-to-text interpreting (termed print interpreting in our earlier studies) is needed for hard-of-hearing and late-deafened people as a communication aid which gives them access to spoken language (cf. Norberg et al. 2015). In a larger project (cf. Tiittula 2009) we investigated the quality of interpretation from various perspectives and with different methods, including eye movement analysis; one of the main questions was how comprehensible the interpretation was and how its methods could be developed in order to enhance accessibility. In the following, we address the reception of speech-to-text interpretation, especially the reading process, through an analysis of eye movements. Our aim is to analyse how the gaze behaviour of the reader reflects the reading process of this kind of dynamic text.

The present work is a pilot study, which consists of an experiment where the participants read a dynamic speech-to-text interpreted text presented on a computer monitor. The text was the output of a speech-to-text interpreting process in which a professional speech-to-text interpreter transformed a spoken conference-like presentation into written format. In Finland, speech-to-text interpreters use a standard QWERTY keyboard; that is, a stenotype is not used. Since speaking is much faster than writing, in spite of a high typing rate, errors and omissions occur. The text appears on the screen letter by letter as the speech-to-text interpreter writes, and recipients see the real-time writing process, including pauses and corrections.

Although the processes of reading and visual perception, in particular, are well studied (cf. Rayner 1998), the research and models are predominantly based on reading static text, often restricted to just one sentence at a time. Today, there is an ever increasing range of dynamic texts enabling different types of reading (for reading web pages, see, for example, Simola 2011), presumably involving different comprehension processes as well. The various formats for presenting dynamic text on screen include scrolling, paging, leading, and RSVP (*Rapid Serial Visual Presentation*). Scrolling involves sliding text displayed, for example, letter by letter or word by word; in Finnish speech-to-text interpreting, both methods are used. Paging presents the text divided into pages that fit the screen. The reader can move one page at a time to continue reading. In leading, dynamic text is scrolled horizontally from right to left on a single line across the screen. Finally, RSVP presents the text in successive chunks of one or more words at a time in a predetermined location on the screen. Consecutive chunks of text are presented at a predetermined rate that may be selected by the user (Muter 1996; Potter 1984). As the use of dynamic texts is growing rapidly in different forms of translation, media presentations and computer-mediated communication, it is important to study the reading process of dynamic texts in order to make their presentation more effective.

In the present study, readers' gaze behaviour is treated from a linguistic perspective. More precisely, the article focuses on regressions of gaze (i.e. backward-directed eye movements) as well as on their relationship to linguistic factors, such as the class of words on which the gaze lands. The aim of the article is to show that the test subjects need to read certain words again in order to construct the meaning of the text. The hypothesis is that these words are mainly content words, most often nouns[1]. The fact that the text is dynamic makes the experiment more interesting: indeed, as the test subjects know that they will have only a limited time to read the text before it disappears, the reading process has to be particularly efficient.

The study consists of quantitative and qualitative analysis. The quantitative analysis deals with regressive eye movements on top of the text being read. Here the number of reread words is examined, with the emphasis on the first and second landing points of regression. The qualitative analysis further scrutinized the examination from a linguistic perspective. More precisely, the qualitative analysis focused on the class of words involved in the first and the second landing points of regression, as well as on the structure and the meaning of the sentences including regressions. Before presenting our study, we will briefly explain the basic terms and findings from previous research into eye movements in reading process (Section 2). In Section 3, we shift the focus to linguistic elements. Section 4 describes the data and the method. The main section (5) consists of the quantitative and qualitative analysis and the results of the eye movement experiment. Section 6 concludes the paper.

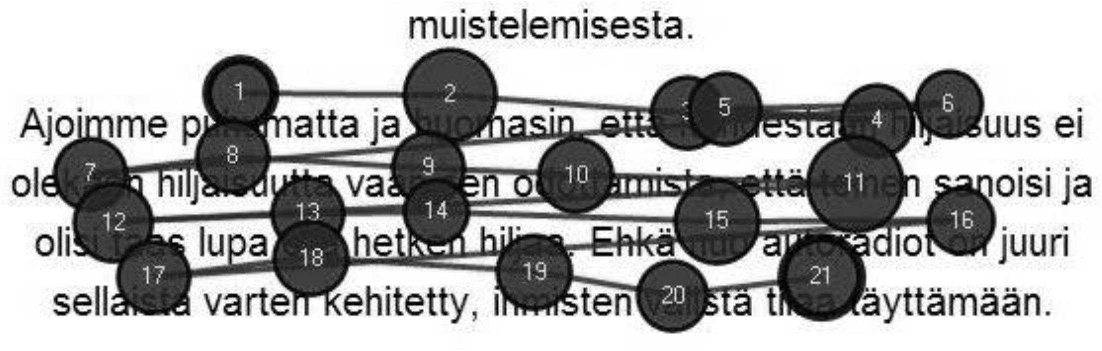
---

[1] Nouns, verbs and adjectives are considered to be 'content words' in this study. All the other words are categorized as 'function words'.

2 Eye movements in the reading process

Reading involves visual processing of the words included in a text. Eye movements in reading can provide a window into the cognitive process of perception and comprehension that takes place during reading. Eye tracking provides eye movement data which demonstrate gaze behaviour in reading. A corneal reflection by infrared light can be recorded through an infrared camera attached to the eye tracker. While reading, eyes make brief jumps along the line of the text, stop for a while and again continue moving forward. Sometimes eyes also move backward. Rapid eye movements from one place to another are known as saccades, while the pauses between saccades are termed fixations. Typical fixation duration in reading an English text is 200–250 ms, and the average saccade length is 7–9 characters (Rayner 1998: 375). Figure 1 presents a typical gaze path in reading with fixations and saccades. The circles marked with numbers are fixations and the straight lines between the circles are saccades.

Figure 1. Gaze path in reading



Eyes do not move steadily forward when reading; rather, they also move backwards for rereading. Eye movements opposite to the direction of written text (right-to-left in the cases where text is read from left-to-right) along the line or movements back to previously read words and lines are called regressions. For the particular time frame in Figure 1, fixations moved forward from number one to four and then moved backward to number five before moving forward again to number six and then on to the next line.

Rereading is a natural human eye movement behaviour in reading. It can indicate an active process that serves a useful function, such as allowing readers to improve text comprehension or fill in gaps in memory about the content of the text (e.g. Levy et al. 1992). Past research has also shown that look-backs or rereadings are often an indicator of comprehension difficulties (Rayner 1998). If the comprehension process does not proceed smoothly, readers tend to look back more. On the other hand, look-back fixations to the most important segments of the text are strategic in nature (Hyönä et al. 2002; Hyönä & Nurminen 2006). Therefore, different eye movements in reading, such as looking at the text for a long time and producing longer fixations, or looking back and rereading, could be caused by different cognitive mechanisms. Short within-word regressions may occur when the reader has difficulty processing the currently fixated word (Carpenter & Just 1983; Rayner & Duffy 1988). Longer regressions that is regressions longer than 10 characters back along the line or to another line, may occur because the reader has failed to understand the text (Carpenter & Just 1983; Rayner & Duffy 1988). Since eye movements reflect difficulties in understanding the document being read, they can also be used to automatically recognize the quality of the text by integrating gaze data from several readers (Biedert et al. 2012).

The lines of text readers look at can be divided into three regions: the foveal region (the central 2º of vision), the parafoveal region (which extends from the foveal region to about 5º on either side of fixation), and the peripheral region (which includes everything beyond the parafoveal region) (Rayner & Pollatsek 2013: 443). Rayner et al. (2006) study demonstrated the importance of preprocessing the word to the right of fixation for fluent reading: When the next word disappeared or was masked, reading was disrupted, which indicates that readers also acquire information from the parafoveal region. Consequently, different types of text representation may affect the viewing pattern. For instance, when text is displayed in scrolling mode, the absence of the word to the right of fixation may cause regressions. Romero-Fresco (2010), who has studied re-spoken live subtitles, compared scrolling mode (word-for-word) presentation to block subtitles. He found that word-for-word subtitles caused almost twice as many fixations as block subtitles, and very often the gaze went back to previous words (ibid. 187–189). Studies by Sharmin et al. (2015) and Sharmin & Wiklund (2014) found more rereading gaze behaviour in reading dynamic text using word-by-word presentation format compared to letter-by-letter format. Different study by Sharmin et al. (2012) found a larger number of regressions and longer fixation duration in reading text presented in small pieces or chunks (30 characters in length) compared to larger pieces of text (sentences or paragraphs), clearly indicating that small textual chunks are more difficult to cognitively process.

In another reading context, while reading for translation, translation students were observed to read static source text differently from dynamic target text that they were producing (Sharmin et al. 2008). Average fixation duration was higher on the dynamic text than on the static source text.

Readers do not, however, fixate on all the words in a text. In particular, many short words are skipped over (Weger & Inhoff 2006). Consequently, foveal fixation on each word is not necessary. Previous studies have shown that content words are fixated on much more than function words (Carpenter & Just 1983; Rayner & Duffy 1988). This can be explained by the frequency and the length of the words and by parafoveal processing (Rayner & Duffy 1986; Staub & Rayner 2007). Consequently, re-fixations are also more likely to target longer words (cf. Vitu et al. 1990). Function words provide information about sentence structure and can be neglected as soon as the structure is clear (Müsseler 2003: 604).

3 A linguistic perspective on the reading process

Discourse processing is often considered to consist of three levels: 1) the surface level, 2) the propositional or textbase level (construction of propositions and their relationships), and 3) the situation or mental model level (van Dijk & Kintsch 1983). The surface level contains only the form of what has been read. The textbase level involves understanding the underlying meanings of what has been read. Finally, at the situation model level, the reader connects the information that has been read with prior knowledge in order to build inferences (van Dijk & Kintsch 1983). In our speech-to-text interpreting data, the regressions seem to be related to processing at the textbase level. That is, the regressions show which items the reader needs in order to construct a coherent meaning (cf. also Kintsch 1988; 1998). This is because, on the one hand, our readers did not only see the form of what had been read (the 'surface level'), but they were also reading in order to understand the meaning of the text (the 'propositional or textbase level'), and on the other, connecting information that has been read with prior knowledge (the 'situation model level') is a cognitive process that is not reflected by eye movements.

According to the Competition Model, a theory of cross-linguistic sentence processing, people take into account several cues contained in the sentence context (MacWhinney & Bates 1989). These cues include such features as word order, morphology, and semantic characteristics, and are used to compute a probabilistic value for all possible interpretations. Then the interpretation that has the highest likelihood is chosen (MacWhinney & Bates 1989). Our hypothesis is that the gaze behaviour of the reader shows which cues are relevant for understanding the meaning of the text.

The purpose of the present study is qualitative in nature: that is, through analysing regression clusters, we aim to clarify the reading process of dynamic text in speech-to-text interpreting. In addition, we examine the parts of speech of the first and the second landing points of regressions during the reading of dynamic text in letter-by-letter presentation format. The first landing point is the first word on which the regression lands. The second landing point is the second word on which the regression lands – that is, it is the landing point of the saccade subsequent to the first regression. The second landing point can be both forward and backward in terms of reading direction. These were not separated in the quantitative analyses, but they were taken into account in the qualitative analyses.

4 Data and methods

4.1 Data

In the experiment, the participants read a dynamic text on a standard computer monitor. The text was a short extract from the outcome of a speech-to-text interpreting process in which a professional speech-to-text interpreter transformed a spoken conference-like presentation in Finnish into written format using the speech-to-text interpreting tool Sprintanium (Špakov 2011). The interpretation was first produced in a live situation, and afterwards rendered at a real-time pace on the screen in letter-by-letter format. The input of the interpretation was a prepared talk, and although it was freely spoken, the clause structures of the text reflect written language. Furthermore, speech-to-text interpretation tends to be more formal than the oral source text, even if the written output appearing on the screen is supposed to correspond to the spoken input as closely as possible, (for example, pronouns and particles tend to be omitted, see Tiittula 2006). Since the experiment focused on reading, the video of the original talk was not shown.[2]

[2] The setting resembles a remote interpreting situation.

Ten participants took part in the experiment, all of whom had normal or corrected-to-normal vision. The average age of the participants was 29.6 years, with a standard deviation (SD) of 11.14 and an age range of 21–51 years. They were all either members of the university staff or students; they were not clients of speech-to-text interpreting.[3] They reported computer use of an average of 5.6 hours per day with a SD of 2.27 and a range of 2–8 hours.

A Tobii T60 remote eye-tracking device was used to track the participants' gaze on its integrated 17-inch TFT colour monitor (with 1280 x 1024 pixels resolution). Tobii Studio eye-tracking analysis software was used to collect the eye movement data. We also used the software for the observational analysis of eye movements.

4.2 Procedure and design

At the beginning of the experiment, all the participants were informed about the test procedure and a background questionnaire was carried out in order to collect information on the participants' education, age, and work experience, etc. We also provided a post-test questionnaire including some questions related to the text and the user's experience of reading the text. The participants were informed about the post-test questionnaire at the beginning of the test in order to motivate them to read the text carefully.

The eye tracker was then calibrated for each participant's eyes. The distance between the eye-tracking monitor, where the text appeared, and the participant was about 60 cm. The stimulus consisted of one short text, which can be seen in Figure 2, at the end of the next section. The total number of words in the test text was 156. The number of content words was 89 (57.1%), and the number of function words 67 (42.9%). The number of nouns was 46 (29.5%).

[3] Because the clients of speech-to-text interpreting are late deafened or hard-of-hearing, many of them are older people. In order to get reliable eye movement data, the current study investigated able bodied participants. However, we conducted a follow-up experiment with a group of hard-of-hearing subjects. There, we found it challenging to obtain eye movement data from the eye tracking system for participants over 65 years of age with thick extraocular muscles. For example, one participant had to push back the extraocular muscles with his fingers to allow the eyes to be tracked.

4.3 Eye movement data and regression landing points

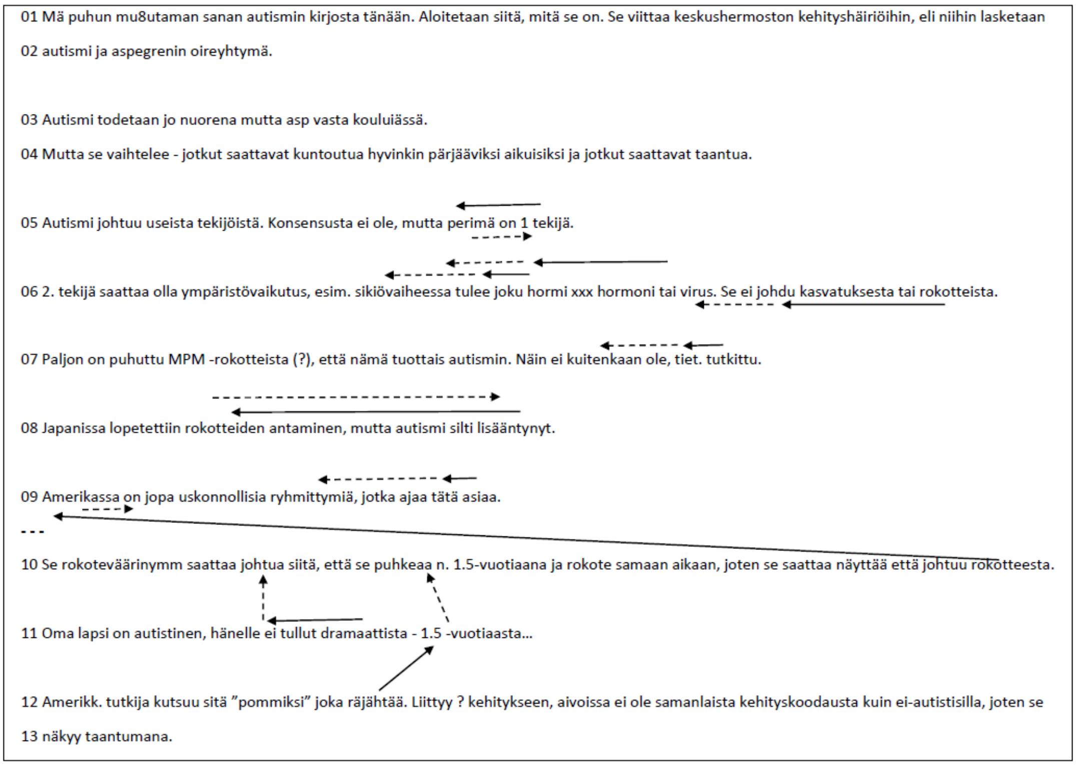
The eye movements of the test participants were detected using an eye-tracking device. Slow motion video recordings of the eye movement data were carefully examined to detect the fixation points and the regression landing points of the words. In reading, the properties of the words fixated on (e.g. word frequency, word length) have been found to influence fixation duration (Liversedge et al. 2011). Prolonged fixations usually indicate more demanding cognitive processing. However, in the present study we were unable to use fixation data in our analysis. A typical velocity-based fixation algorithm considers two gaze points as belonging to the same fixation if their distance is below a specified threshold value. In the reading of static text, gaze jumps ahead in steps of typically 7–9 characters (Rayner 1998: 375), which makes it easy to distinguish fixations from each other. However, the reading process is fundamentally different in our study, in which dynamic speech-to-text interpreted text was presented in letter-by-letter format where the gaze followed the gradual appearance of letters on the screen. In such a context, the two consecutive data points produced by the eye tracker are relatively close together, making it impossible to distinguish fixations from one another. Indeed, one may question whether reading in this case is based on usual fixations at all, as it more closely resembles the smooth pursuit of emerging text (Räihä et al. 2011).

Kruger & Steyn (2013) have also noted the problem of using the standard fixation-based analysis in connection with dynamic text, and they suggest a new metric, a reading index for dynamic texts in subtitling, assuming that gaze data can be reliably classified into fixations and saccades. However, this was also unsuitable for our purposes. Therefore, we adopted a different approach, choosing to consider the number of regressions as an eye movement metric. Regressions have been used in a number of eye-movement studies. For instance, Sanders & Stern (1980) also used regression to study the effects of text characteristics, and Ashby et al. (2005) found that regression reflected the reading proficiency of readers. Furthermore, for dynamic text Specker (2008) used regression as an additional metric to support the fixation-based analysis of eye movements in subtitles.

On the basis of the video observation of eye movements, we marked several regression landing points. In order to avoid complicated and densely populated fixation points, we reprinted the landing points on the stimuli as shown in Figure 2, with the first two regression landing points as the maximum in a row. Figure 2 thus presents some examples of the regression landing points in our data. The arrows with a solid line are used to indicate the first landing points, and the arrows with a dotted line are used to indicate the second landing points.[4] After that, we classified the first and second landing points of the regressions of gaze according to the part of speech of the word in question.

As mentioned earlier, regressions are points where the eye movements of the subject show that s/he is rereading a part of the text. Regressions may occur when the speech-to-text interpreter pauses and no new text is visible, allowing the reader time to check the overall meaning of the text s/he has read. Nevertheless, regressive eye movements also occur in places without pauses, thereby indicating processing difficulties. In these cases, we can assume that the reader regresses until the problem is solved. These different situations (that is, whether or not there was a pause) were not separated in the quantitative analyses, but they were taken into account in the qualitative analyses.

Figure 2. Examples of regression landing points



[4] The line numbers were added to the pictures afterwards in order to facilitate reference to the examples in the text.

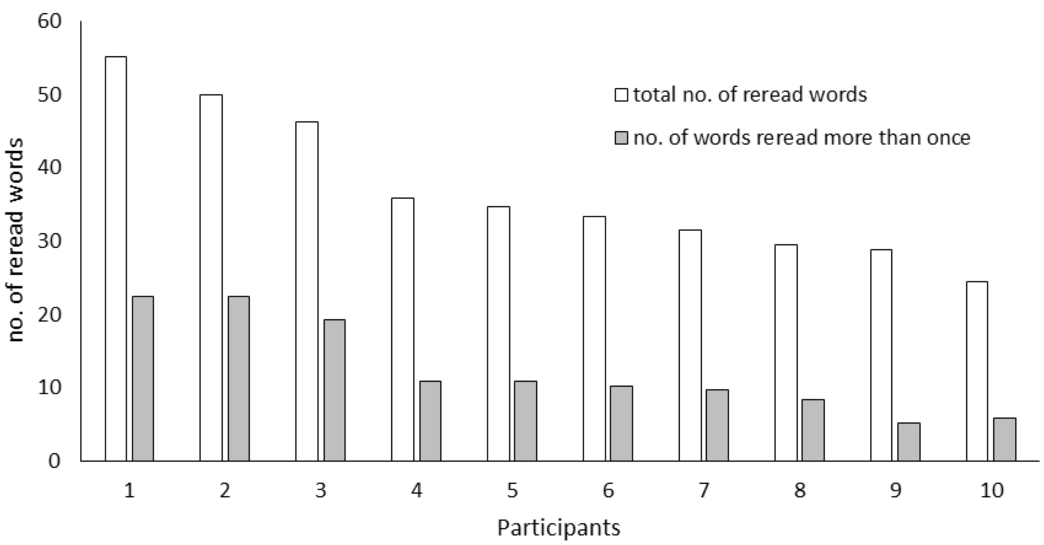
5 Analysis and results

Eye movement videos from each participant for the first 4.5 minutes of reading were used in the analysis, this being the point at which the page began to scroll to allow continued reading. As mentioned earlier, a careful observational analysis was carried out in order to identify both the words from which the regressions or rereading started and, at the same time, the words on which the regressions landed.

5.1 Eye movements in reading speech-to-text interpreted text

All the participants read speech-to-text interpreted text that appeared letter-by-letter. As we were unable to analyse the data by means of fixation duration or fixation count (typical eye movement metrics), we continued our investigation by concentrating on the number of regressions or rereadings by each individual participant. On the basis of the regression landing points, we calculated the number of reread words on the text. Figure 3 presents the total number of reread words for each participant along with the number of words reread more than once. As can be seen, we found variation in reading among the participants in respect to the number of reread words, with participant 1 rereading almost twice as much as participant 10.

Figure 3. The number of reread words for each participant



Our stimuli consisted of speech-to-text interpreted text containing abbreviated words and spelling errors. We found that the frequency of rereading misspelled or abbreviated words was significantly higher than the frequency of rereading the other words in the text. A paired samples t-test produced a statistically significant result in this regard, with $p < .01$, $t = 3.891$ and $df = 9$. Mistyped and abbreviated words are, however, a heterogeneous group. For example, established abbreviations (e.g. *esim.* for *esimerkiksi* 'for instance') are presumably easier to process than the complete word, whereas ad-hoc abbreviations may cause more difficulty. Nevertheless, in order to investigate this issue in greater depth, we need more data.

On the other hand, we also observed a high density of regressions beginning at the end of sentences. A paired samples t-test showed that the number of regressions starting from the last word of a sentence was significantly higher than the number of regressions starting from words in other positions in the sentence ($p < .001$, $t = 13.56$ and $df = 9$).

5.2 Analysing regression landing points

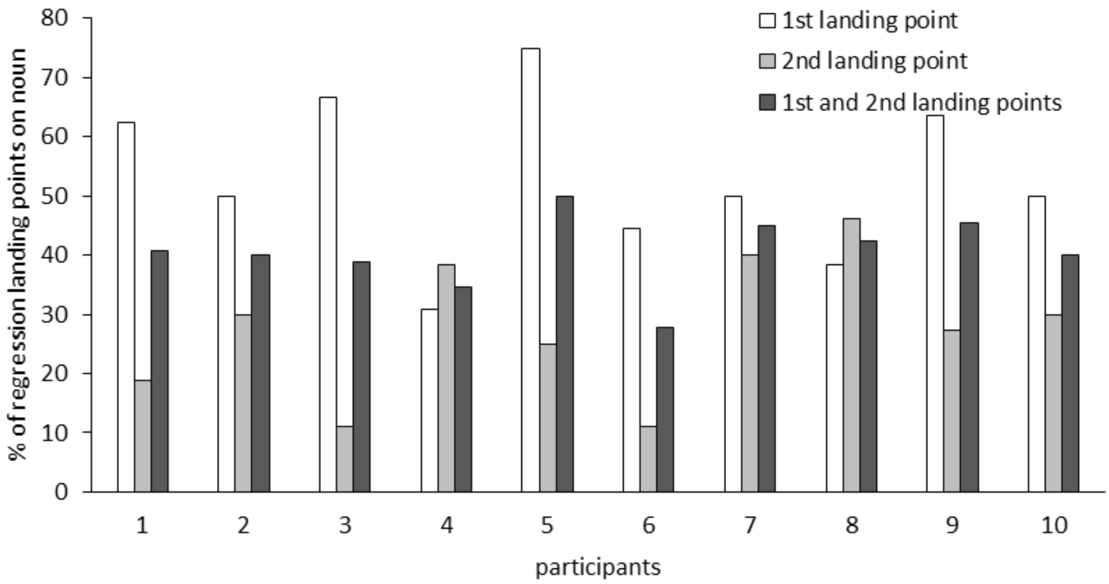
As mentioned earlier, we examined the words at the first two regression points, with our data consisting of a total of 109 regression clusters. All ten informants were represented in the regression cluster data, as well as in all the other figures provided in this article. Analysis revealed individual differences between participants in the number of regressions. The number of regression clusters per participant varied from 8 to 16, with the average number of regression clusters per participant being 10.9.

We analysed the first and second landing points of the regression clusters, with the sum of the first and the second landing points being 218 (= 109 + 109), by sorting them according to parts of speech.

When we considered both the first and the second landing points of the regressions together, we found that the first or the second landing point was a noun in 74.3% (81) of cases, even though only 29.5% (46) of all words in the data were nouns.

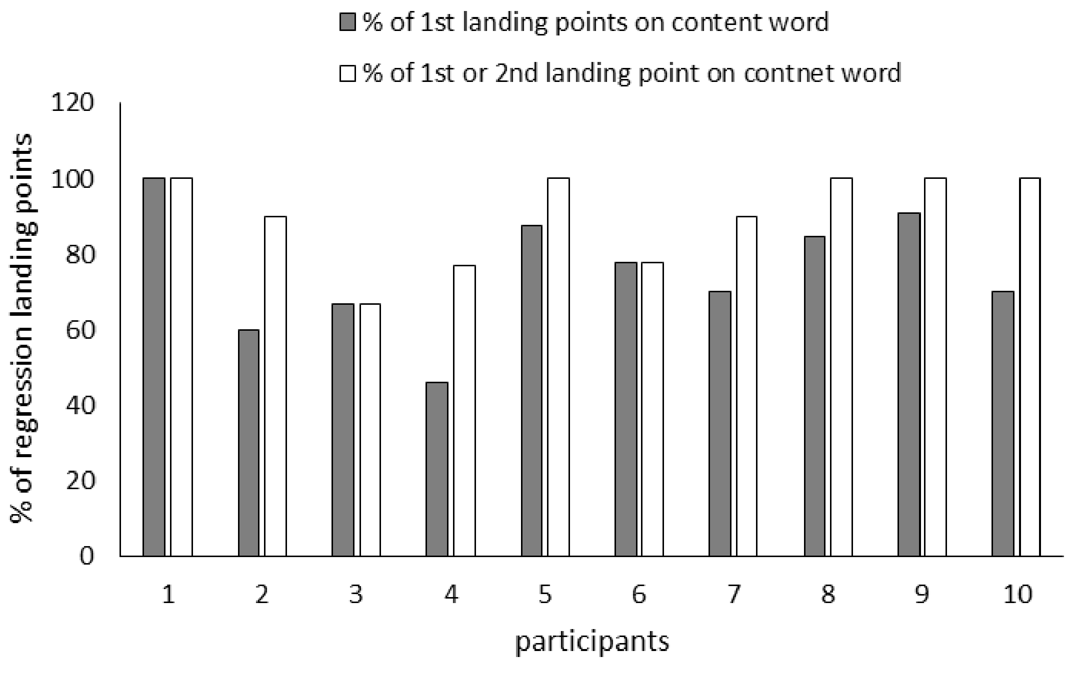
Figure 4 below presents the percentage distributions for first and second landing points involving nouns for all the participants.

Figure 4. Percentages of regression landing points involving nouns



When the first or second landing point was not a noun, it was most likely to be a verb or an adjective. In 90.8% (99) of cases, the first or second landing point was a content word. This is interesting because only 57.1% (89) of the words in the data were content words. Figure 5 presents the distribution of first two regression landing points on content words for all the participants.

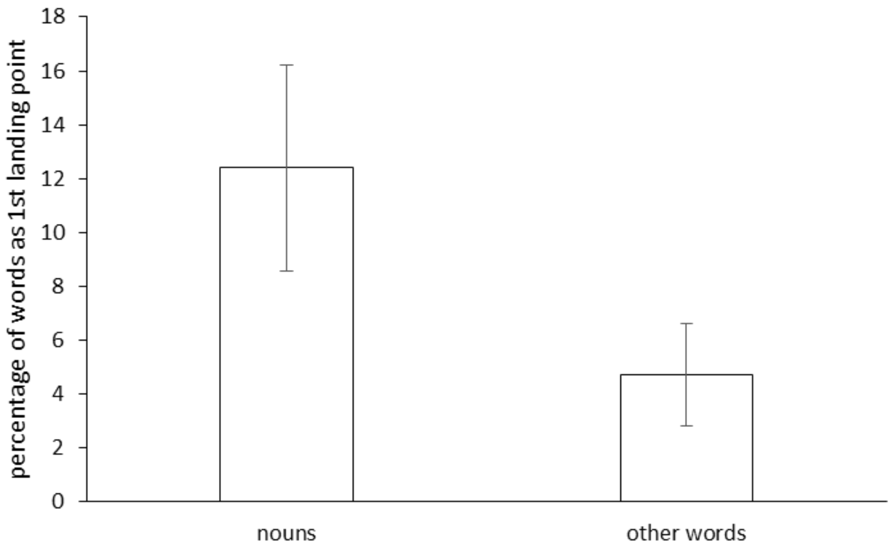
Figure 5. Percentages of regression landing points on content words



Function words, in turn, were rarely regression landing points: the first or the second landing point was a function word in only 9.2% (10) of cases, even though 42.9% (67) of the words in the data were function words.

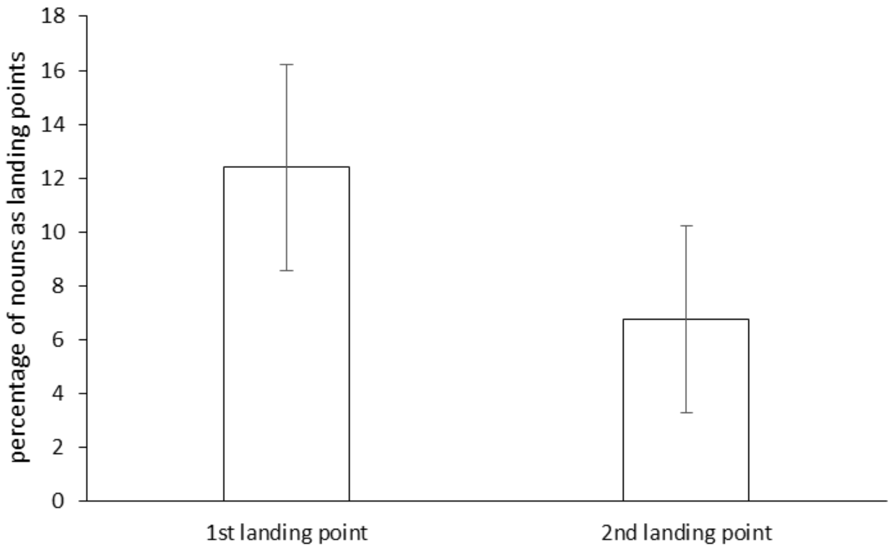
Although less than 30% of the words in the text were nouns, a one-way ANOVA found that the first regression landing point was more likely to occur on a noun than on any other category of words. The finding is significant, with $p < .001$ and $F_{1,19} = 31.937$ (Figure 6). Average percentage number of nouns as a first regression landing point is 12.4 while the number is 4.7 for the other words with variances 14.8 and 3.6 respectively.

Figure 6. Percentage of nouns and other words at the first regression landing points by participant, with standard deviation error bars



In contrast, for the second landing points, nouns were no longer the primary targets, thus demonstrating that in regressive reading the first landing point is more likely to be noun than is the second landing point (Figure 7). This difference was statistically significant ($p < .01$, and $F_{1,19} = 11.929$). Average percentage number of nouns as a second landing point is 6.7 with variance 12. However, if we combine both the first and the second landing points, we still see that nouns are the most significant targeted words as landing points ($p < .001$, and $F_{1,19} = 78.193$).

Figure 7. Percentage of nouns at the first and the second landing points with standard deviation error bars.



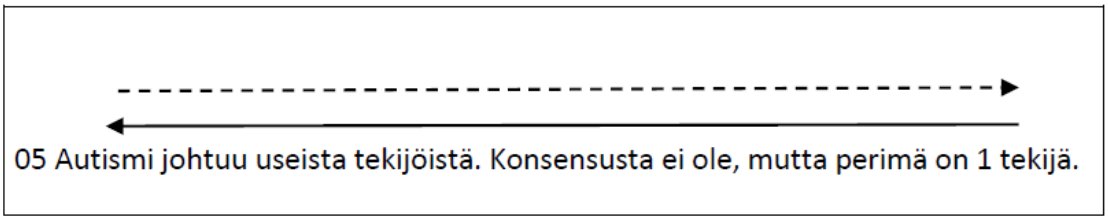
One explanation for the high proportion of nouns as primary landing points could be word length (cf. Vitu et al. 1990). In our data the mean word length of nouns was 9.2 characters, compared to 7.9 characters for other content words and 3.9 characters for function words.

5.3 Linguistic description of regression landing points

We will now present some examples illustrating different situations where regression clusters occur in the whole dataset, including the test results of all ten informants. The situations selected as examples provide a quantitative perspective on the different cases presented in our quantitative analyses. In addition, the aim of the examples is to present a variety of contexts in which regression clusters occur and to find possible explanations for the eye movement. In the linguistic analysis we used both pictures of the regression landing points (cf. Figure 2) and videos showing each participant's eye movements.

In our data, the first landing point of a regression was a noun in 52.3% (57) of cases. For instance, in example 1 (*Autismi johtuu useista tekijöistä. Konsensusta ei ole, mutta perimä on 1 tekijä.* 'Autism is caused by several factors. There is no consensus, but genotype is 1 factor.', line 05), the first landing point was a noun. As the arrow in Figure 8 shows, the gaze first went to the word *autismi* ('autism'), which is a noun. Then it returned to the starting point – that is, to the word *tekijä* ('factor'), which is also a noun.

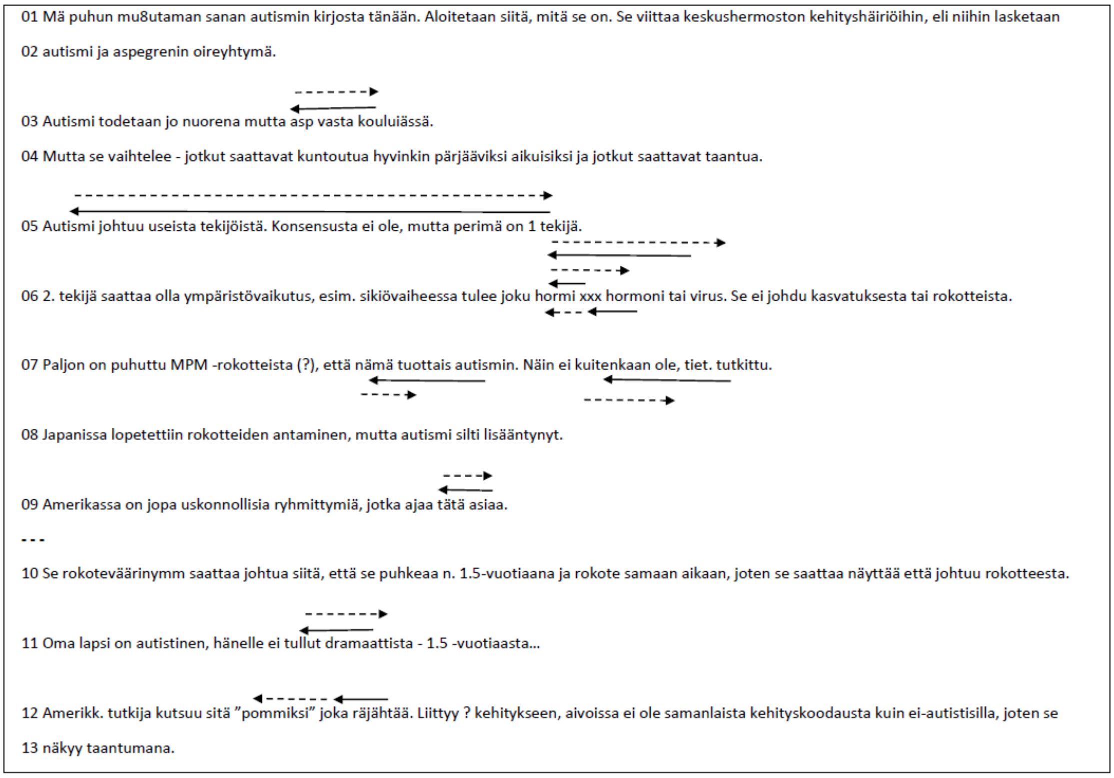
Figure 8. Example 1



The context of Example 1 is given in Figure 9.

After the last word, *tekijä* ‘factor’, a longer pause occurred, and the test subject stopped reading and fixated on ‘factor’. The pause continued, with no new text visible, so the gaze jumped to the beginning of the line, landing on the first word (‘autism’), which is a noun as well as the subject of the sentence and the topic of the whole line. From ‘autism’ the gaze returned to the end of the line, to ‘factor’. Here, the reader seemed to be trying to establish a coherent representation or confirm the overall meaning of the sequence in which the missing link must be inferred between the first and the second sentences (‘Autism is caused by several factors.’ and ‘There is no consensus, but the genotype is 1 factor.’, line 05). The eye movement from the beginning to the end of the line can, however, have another explanation: the reader may have been checking to see whether the text would continue. However, on the video it can be seen that the text did not continue, and the gaze jumped once more to ‘autism’ and from there to the beginning of the next line, where new text began to appear.

Figure 9. The context Example 1


01 Mä puhun mu8utaman sanan autismin kirjosta tänään. Aloitetaan siitä, mitä se on. Se viittaa keskushermoston kehityshäiriöihin, eli niihin lasketaan
02 autismi ja aspegrenin oireyhtymä.
03 Autismi todetaan jo nuorena mutta asp vasta kouluiässä.
04 Mutta se vaihtelee - jotkut saattavat kuntoutua hyvinkin pärjääviksi aikuisiksi ja jotkut saattavat taantua.
05 Autismi johtuu useista tekijöistä. Konsensusta ei ole, mutta perimä on 1 tekijä.
06 2. tekijä saattaa olla ympäristövaikutus, esim. sikiövaiheessa tulee joku hormi xxx hormoni tai virus. Se ei johdu kasvatuksesta tai rokotteista.
07 Paljon on puhuttu MPM -rokotteista (?), että nämä tuottais autismin. Näin ei kuitenkaan ole, tiet. tutkittu.
08 Japanissa lopetettiin rokotteiden antaminen, mutta autismi silti lisääntynyt.
09 Amerikassa on jopa uskonnollisia ryhmittymiä, jotka ajaa tätä asiaa.
---
10 Se rokoteväärinymm saattaa johtua siitä, että se puhkeaa n. 1.5-vuotiaana ja rokote samaan aikaan, joten se saattaa näyttää että johtuu rokotteesta.
11 Oma lapsi on autistinen, hänelle ei tullut dramaattista - 1.5 -vuotiaasta...
12 Amerikk. tutkija kutsuu sitä ”pommiksi” joka räjähtää. Liittyy ? kehitykseen, aivoissa ei ole samanlaista kehityskoodausta kuin ei-autistisilla, joten se
13 näkyy taantumana.

The first landing point was the closest noun in 33.0% (36) of cases. In Example 2 (*Aivoissa ei ole samanlaista kehityskoodausta kuin ei-autistisilla.* ‘In the brain there is not the same kind of developmental coding as in the brain of non-autistic individuals.’, line 12), regression began from the word *samanlaista* (‘the same kind of’), and the first landing point was the word *aivoissa* (‘in the brain’), which is the closest noun. After that, the gaze returned towards the starting point of the regression and then passed it (cf. Figure 10). Next, the gaze stopped at the compound *kehityskoodausta* (‘developmental coding’), which is also a noun.

Figure 10. Example 2

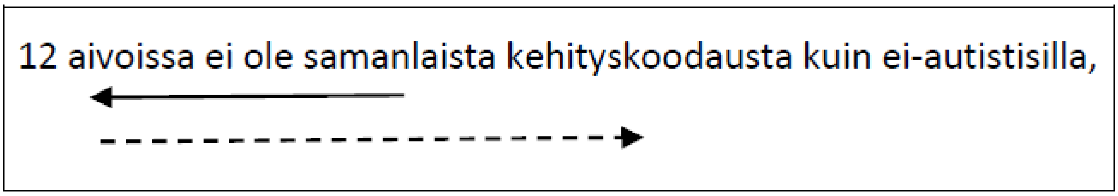
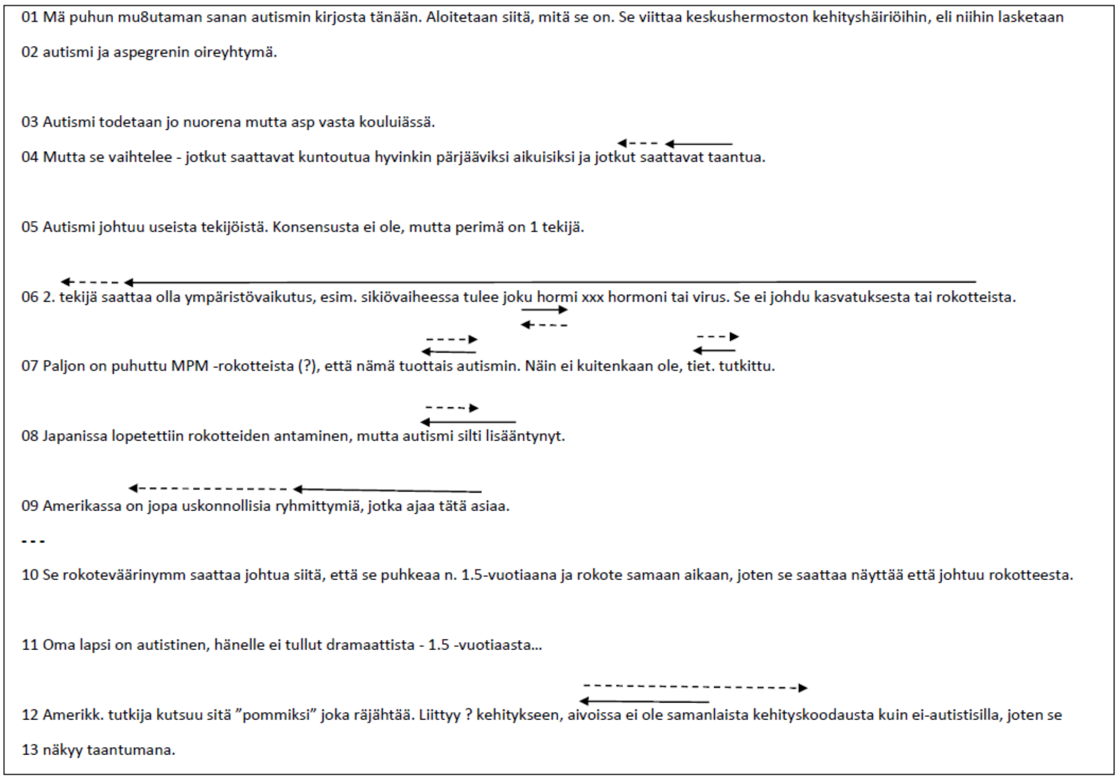

12 aivoissa ei ole samanlaista kehityskoodausta kuin ei-autistisilla,

Figure 11 shows the context of Example 2.

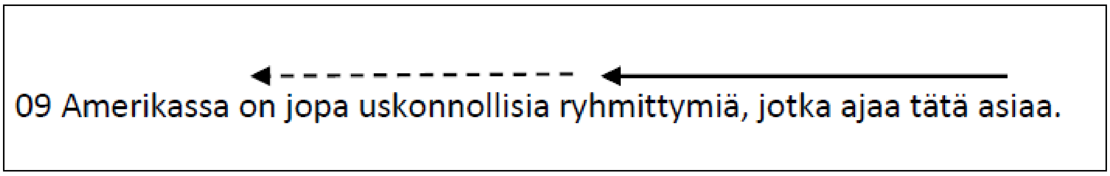
Figure 11. The context Example 2



The word *samanlaista* (‘the same kind’, line 12) needs a head word, which was sought in the nearest preceding noun, *aivoissa* (‘in the brain’, line 12), the landing point of the regression. This noun cannot be the headword because it is in the inessive case, which expresses location. Thus, the gaze returned to the starting point, followed the emerging text, and stopped at the next noun, which is the head word of ‘the same kind’. The difficulty of this sentence is due to its lack of coherence: because of omissions, the previous sentences in the line are neither connected nor linked to the prior text.

In Example 3 (*Amerikassa on jopa uskonnollisia ryhmittymiä, jotka ajaa tätä asiaa.* ‘In America there are even religious groups that champion this cause.’, line 09), the first landing point was the word *ryhmittymiä* (‘groups’), which is also a noun (cf. Figure 12). After the first landing point, the regression continued in the same direction until it reached the auxiliary verb *on* (‘are’), appearing three words earlier.

Figure 12. Example 3



The context of Example 3 is presented above in Figure 11.

At the end of the line, a long pause occurred during which the described eye movements took place. The last words (*tätä asiaa*, ‘this cause/issue’) are anaphoric, referring to something earlier in the text. The regressive eye movement stopped at the closest noun (*ryhmittymiä*, ‘groups’), which is the head word of the relative clause. According to Staub & Rayner (2007), short or frequently occurring words (such as *jotka*, ‘that’, and *ajaa*, literally ‘drive’ in the relative clause) can be skipped, whereas unpredictable words are fixated on for longer. The phrase *uskonnollisia ryhmittymiä* (‘religious groups’) can be regarded as unpredictable in the context of autism, especially because the sentence is not explicitly connected to the prior sentence (‘In Japan, vaccinations were stopped, but autism has still increased.’). From this point the gaze regressed further, landing on the auxiliary verb *on* (‘are’), which is one of the most frequent words and very short; the long fixation on this place can be explained by the linguistic processing of the overall meaning of the sentence. Even though the meaning of the lexical entities have been processed, constructing a coherent discourse representation can be difficult. The video showed that after this the gaze moved further back to the adverbial phrase *Amerikassa* (‘in America’), then returned to the verb of the relative clause (‘champion’), where it stopped. The gaze then continued to the end of the sentence and jumped to the beginning of the new line as soon as new text began to appear.

In Example 4 (*Se viittaa keskushermoston kehityshäiriöihin, eli niihin lasketaan autismi ja aspegrenin*[5] *oireyhtymä*. ‘It indicates developmental disorders of the central nervous system, that is they include autism and Asperger’s syndrome’, lines 01–02), the first landing point was the word *aspegrenin* (‘aspegren’s’), which is the closest noun to the starting point *oireyhtymä* (‘syndrome’) (cf. Figure 13). (The first landing point was also the closest noun in Examples 2 and 3.) After the word *aspegrenin*, the gaze returned to the starting point of the regression.

[5] Here the interpreter makes a mistake: the speaker said Asperger’s, not aspegren’s.

Figure 13. Example 4

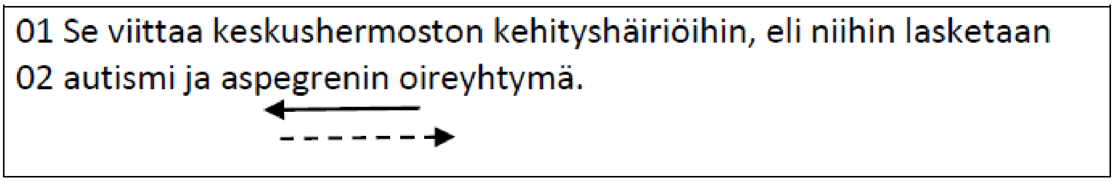
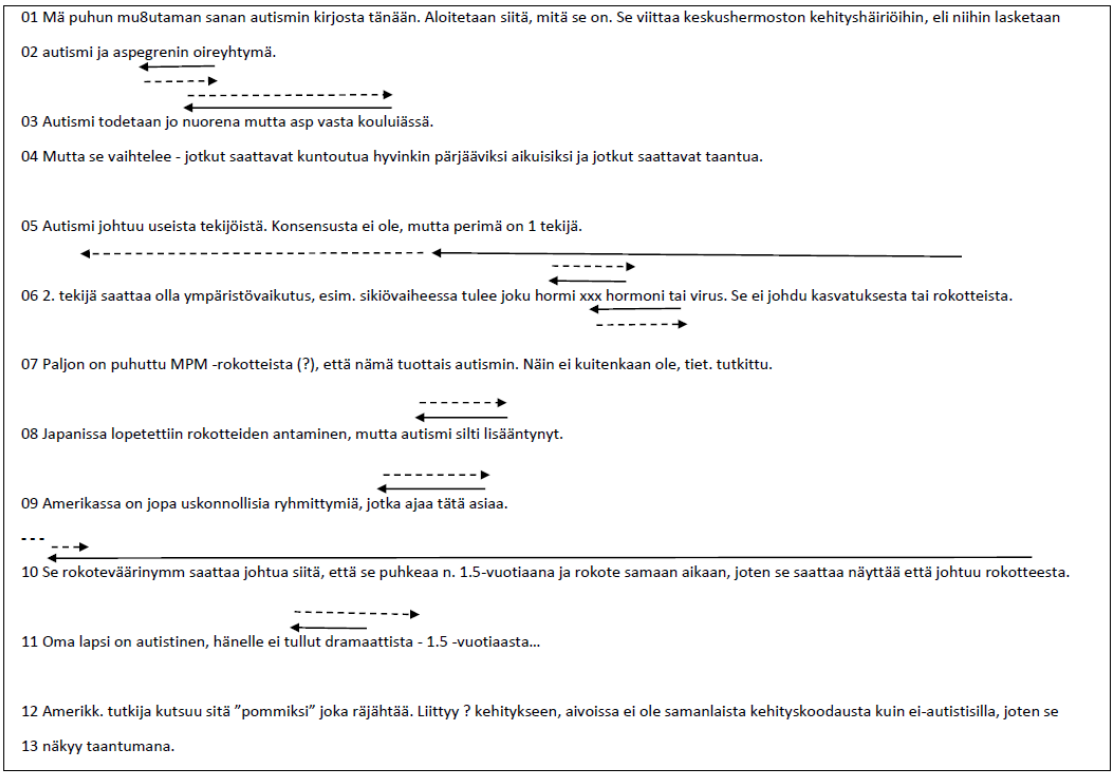

01 Se viittaa keskushermoston kehityshäiriöihin, eli niihin lasketaan
02 autismi ja aspegrenin oireyhtymä.

Figure 14 shows the context of Example 4.

Figure 14. The context Example 4


01 Mä puhun mu8utaman sanan autismin kirjosta tänään. Aloitetaan siitä, mitä se on. Se viittaa keskushermoston kehityshäiriöihin, eli niihin lasketaan
02 autismi ja aspegrenin oireyhtymä.
03 Autismi todetaan jo nuorena mutta asp vasta kouluiässä.
04 Mutta se vaihtelee - jotkut saattavat kuntoutua hyvinkin pärjääviksi aikuisiksi ja jotkut saattavat taantua.
05 Autismi johtuu useista tekijöistä. Konsensusta ei ole, mutta perimä on 1 tekijä.
06 2. tekijä saattaa olla ympäristövaikutus, esim. sikiövaiheessa tulee joku hormi xxx hormoni tai virus. Se ei johdu kasvatuksesta tai rokotteista.
07 Paljon on puhuttu MPM -rokotteista (?), että nämä tuottais autismin. Näin ei kuitenkaan ole, tiet. tutkittu.
08 Japanissa lopetettiin rokotteiden antaminen, mutta autismi silti lisääntynyt.
09 Amerikassa on jopa uskonnollisia ryhmittymiä, jotka ajaa tätä asiaa.
10 Se rokoteväärinymm saattaa johtua siitä, että se puhkeaa n. 1.5-vuotiaana ja rokote samaan aikaan, joten se saattaa näyttää että johtuu rokotteesta.
11 Oma lapsi on autistinen, hänelle ei tullut dramaattista - 1.5 -vuotiaasta...
12 Amerikk. tutkija kutsuu sitä "pommiksi" joka räjähtää. Liittyy ? kehitykseen, aivoissa ei ole samanlaista kehityskoodausta kuin ei-autistisilla, joten se
13 näkyy taantumana.

The object of the last clause consists of a noun ('autism') and a noun phrase (*aspergrenin oireyhtymä*, 'aspergren's syndrome', line 02). 'Autism' is the topic of the talk and part of the theme of the last sentence; accordingly, it should be easy to process, whereas 'aspergren's' seemingly caused difficulties, as it is a medical term, an infrequent word and, additionally, mistyped. The last word, *oireyhtymä* ('syndrome'), is a compound ('symptom' + 'group'). According to Hyönä & Pollatsek (1998; also Pollatsek & Hyönä 2005), the frequency of the first morpheme influences fixation. In our example, the reader seems to have been able to anticipate the end of the compound, since s/he did not wait for its appearance but regressed from *oireyh-* to the attribute ('aspergren's'), a word which must be understood in order to comprehend the whole noun phrase.

The first landing point was a content word in 76.1% (83) of the cases. In the fifth example (*Paljon on puhuttu MPM -rokotteista (?), että nämä tuottais autismin.* 'There has been a lot of discussion about MPM vaccines (?) that they would cause autism.'[6], line 07) (cf. Figure 15), the first landing point was the verb *tuottais* ('would produce', i.e. 'cause'), which is a content word as well as the closest word to the starting point of the regression. This is the case in 43.1% (47) of the regressions.

Figure 15. Example 5

The context of Example 5 is presented in Figure 11.

In Example 5, the regressive eye movement took place from the last word of the clause ('autism') to the previous verb *tuottais* ('would produce'). The subject of the clause is *nämä* ('these'), which refers to vaccines. Vaccine/vaccination less frequently collocates with 'produces' than, say, 'protects',[7] and thus the clause is less predictable, which perhaps explains the regression.

Example 6 (*Amerikk tutkija kutsuu sitä "pommiksi" joka räjähtää*, 'An American researcher calls it a "bomb" that will explode', line 12) presents another instance where the first landing point was the closest content word. Here, the regression started from the word *räjähtää* ('explode') and first landed on the noun "*pommiksi*" ('a "bomb"'), which is the closest content word and also the closest noun. The second landing point was the verb *kutsuu* ('calls'), appearing two words earlier (cf. Figure 16).

[6] This sentence also contains a small mistake by the interpreter, because the speaker is obviously talking about MPR vaccines and not about MPM vaccines.

[7] E.g., a simple Google search produces 7860 hits for the collocation *rokote suojaa* 'vaccination protects', 1120 hits for *rokote aiheuttaa* 'causes', and only 137 hits for *rokote tuottaa* 'produces'.

Figure 16. Example 6

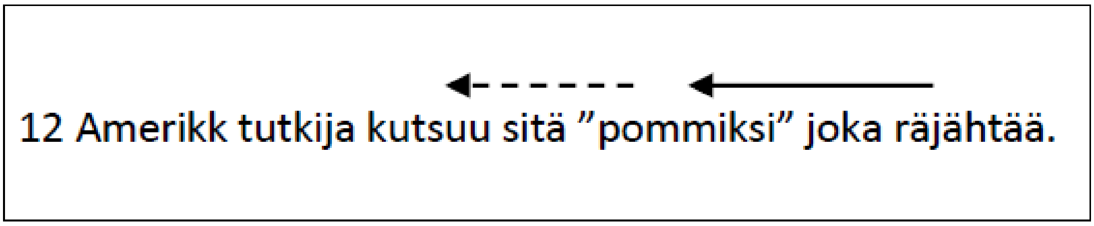
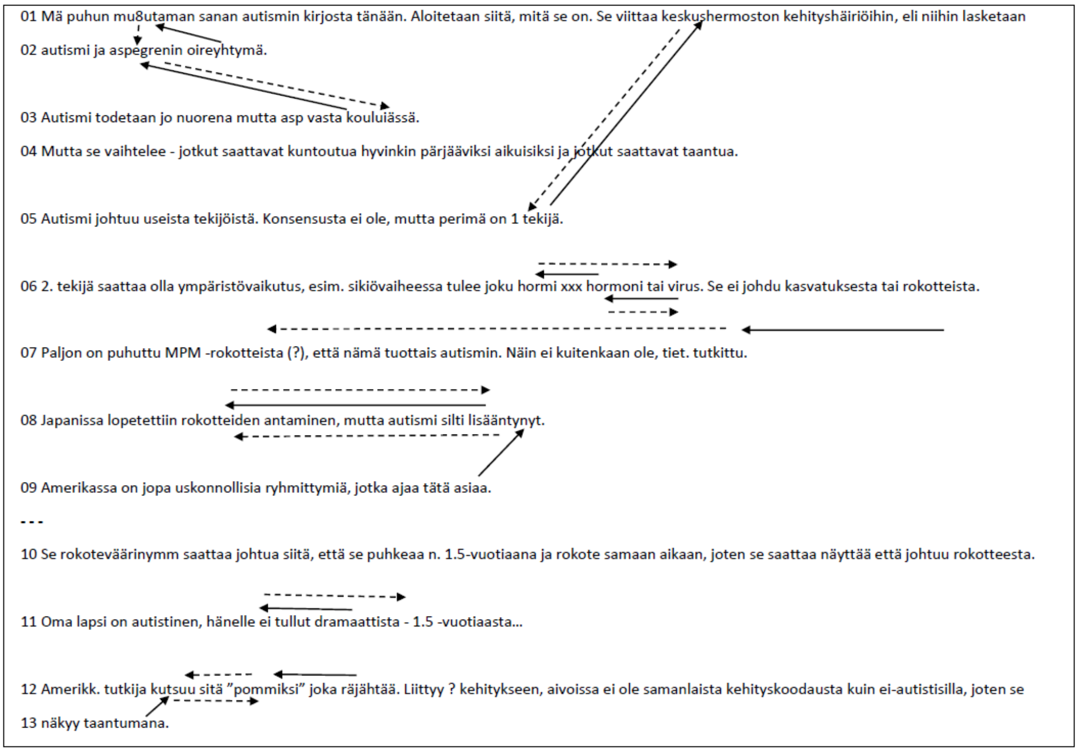


Figure 17 presents the context of Example 6. Prior to the regression, the relative pronoun *joka* (‘that’) was misprinted and corrected, causing a longer fixation on the clause. The end of the predicate was not yet visible (but was presumably predictable) when the regression started. It landed on the head word *pommiksi* (‘a bomb’), the closest content word. Though the noun is in quotation marks, indicating metaphoric use, constructing a coherent discourse meaning can still be difficult, since due to omissions the sentence was not connected with the previous text. From the noun, the regression continued to the verb *kutsuu* (‘calls’), which confirms the metaphoric use of ‘bomb’. At this point, the processing of the sentence seemed to be complete, since, as can be seen on the video, the gaze moved to the beginning of the next sentence.

Figure 17. The context Example 6



6 General findings and conclusions

Rayner (1998: 375) observed that content words are fixated on more often (85% of the time) than function words. Our present analytical approach, from a linguistic perspective, supports these findings. A study by Sharmin et al. (2015) found three different types of gaze behaviour among the participants in reading dynamic text on the basis of fixation frequency and rereading. Although there exists variation in reading behaviour, we found consistency in regression landing points. We analysed fixated words in rereading, hence regression landing points. Our results demonstrated that in reading speech-to-text interpreted text, the first and second landing points in regressions are mostly content words (in 90.8% cases), even though the proportion of content words in the whole data set was only 57.1%. Therefore, the same lexical hierarchy that concerns fixations seems to apply to the landing points of regressions.

O'Grady's findings on language acquisition (1987) suggest that nouns are 'primary' elements of language because they are characterized by autonomous meaning and function. Their referents are perceptively distinct and coherent, whereas verbs, for instance, have a more fragmented meaning. The referents of verbs are not 'present' in the perceptive field as concretely as the referents of nouns (Gentner 1982; Maratsos 1991; Caselli et al. 1995). Therefore, O'Grady (1987) considers verbs and adjectives to be 'secondary' elements. They depend on a relationship to at least one primary element. Function words, in turn, are 'tertiary' elements, because they depend on a relationship to at least one secondary element. The same lexical hierarchy that characterizes language acquisition seems to apply to our data. Thus, our test subjects sought primary and secondary elements of language in order to construct the meaning of what they had just read. Nouns, which are primary elements, were the most likely landing points of regressions.[8] Nouns were, however, longer words than other content words and function words, which may also have influenced the results. Nevertheless, the landing points of regressions reflect the cognitive processing of language through which meaning is constructed.

Speech-to-text interpreted text is dynamic text which can be displayed through scrolling, as was the case in our experiment. Regressions typically occur during pauses, and pauses enable rereading. Speech-to-text

[8] Furtner's & Sachse's (2007) results are similar to ours concerning the importance of the noun for the improvement of text comprehension.

interpreted text is also a special kind of dynamic text because (depending on the methods and tools) omissions and typing errors frequently occur. Our pilot study shows how this affects reading and impedes processing of the text. Omissions lead to incoherence; in many cases, the connection of an element to the previous text was unclear. Accordingly, lack of coherence seems to be a typical cause of regressions. Further possible causes are difficult concepts, infrequent words and unpredictable collocations. Misspellings and ad-hoc abbreviations may also cause regressions and slow the speed of text processing.

Due to the difference between speech rate and writing speed, speech-to-text interpreters are under the constant pressure of time. Accordingly, condensation of the message is necessary. Consequently, many words are omitted or if not omitted, the interpreter risks finishing sentences with increasing delay. According to our earlier research, omitted words are often particles, connectors and pronouns, that is, omissions tend to follow the lexical hierarchy. The results of our pilot study indicate that this can be a reasonable condensation strategy. By mainly reducing tertiary language elements, the interpreter can speed up typing and try to render the primary elements.

The results may also have implications for other types of dynamic texts used in different forms of translation, media presentations and computer-mediated communication. Consciousness of lexical hierarchy may help text producers improve the presentation of dynamic texts or condense the spoken message. The presentation of dynamic texts could be improved, for example, by highlighting the primary elements and by only reducing tertiary language elements.

Acknowledgements

The research was conducted in the project "SpeechText. Print interpreting: process, comprehensibility, and technology" funded by the Academy of Finland. We also wish to thank Helsinki Collegium for Advanced Studies and TraST: Translation Studies and Terminology Research Community for the financial support, as well as University of Tampere for the gaze lab facility.

References

Ashby, Jane & Rayner, Keith & Clifton Jr., Charles. 2005. Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *The Quarterly Journal of Experimental Psychology* 58(6). 1065–1086.

Biedert, Ralf & Dengel, Andreas & Elshamy, Mostafa & Buscher, Georg. 2012. Towards robust gaze-based objective quality measures for text. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, 201–204. New York: ACM.

Carpenter, Patricia A. & Just, Marcel Adam. 1983. What your eyes do while your mind is reading. In Rayner, Keith (ed.), *Eye movements in reading: Perceptual and language processes*, 275–307. New York: Academic Press.

Caselli, Maria Cristina & Bates, Elizabeth & Casadio, Paola & Fenson, Judi & Fenson, Larry & Sanderl, Lisa & Weir, Judy. 1995. A crosslinguistic study of early lexical development. *Cognitive Development* 10. 159–199.

Furtner, Marco & Sachse, Pierre. 2007. Eye movements and text understanding: New evidence for the role of the parts of speech and the regressive fixations. (Paper presented at the conference *ECEM2007 – 14th European Conference on Eye Movements*, Potsdam, August 19–23, 2007.)

Gentner, Dedre. 1982. Why nouns are learnt before verbs: Linguistic relativity versus natural partitioning. In Kuczaj, Stan (ed.), *Language development*, vol. 2, 67–88. Hillsdale: Erlbaum.

Hyönä, Jukka & Nurminen, Anna-Mari. 2006. Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology* 97. 31–50.

Hyönä, Jukka & Lorch, Robert F. & Kaakinen, Johanna K. 2002. Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology* 94. 44–55.

Hyönä, Jukka & Pollatsek, Alexander. 1998. Reading Finnish compound words: Eye fixations are affected by component morphemes. *Journal of Experimental Psychology: Human Perception and Performance* 24. 1612–1627.

Kintsch, Walter. 1988. The use of knowledge in discourse processing: A construction-integration model. *Psychological Review* 95. 163–182.

—— 1998. *Comprehension: A paradigm for cognition.* Cambridge, UK: Cambridge University Press.

Kruger, Jan-Louis & Steyn, Faans. 2013. Subtitles and eye tracking: Reading and performance. *Reading Research Quarterly* 49(1). 105–120.

Levy, Betty A. & Dipersio, Robert & Hollingshead, Ann. 1992. Fluent rereading: Repetition, automaticity, and discrepancy. *Journal of Experimental Psychology: Learning, Memory and Cognition* 18. 957–971.

Liversedge, Simon P. & Gilchrist, Iain & Everling, Stefan (eds.). 2011. *The Oxford handbook of eye movements.* New York: Oxford University Press.

MacWhinney, Brian & Bates, Elizabeth. 1989. *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.

Maratsos, Michael. 1991. How the acquisition of nouns may be different of that of verbs. In Krasnegor, Norman A. & Rumbaugh, Duane M. & Schiefelbusch, Richard L. & Studdert-Kennedy, Michael (eds.), *Biological and behavioral determinants of language development*, 67–88. Hillsdale: Erlbaum.

Müsseler, Jochen. 2003. Periphere und zentrale Prozesse beim Lesen [Peripheral and central processes in reading]. In Rickheit, Gert & Herrmann, Theo & Deutsch, Werner (eds.), *Psycholinguistik. Ein internationales Handbuch* [Psycholinguistics. An international handbook]. (Handbücher zur Sprach- und Kommunikationswissenschaft [Handbooks of linguistics and communication science]) (HSK) 24, 600–608. Berlin: De Gruyter Mouton.

Muter, Paul. 1996. Interface design and optimization of reading of continuous text. In van Oostendorp, Herre & de Mul, Sjaak (eds.), *Cognitive aspects of electronic text processing*, 161–180. Norwood: Ablex.

Norberg, Ulf & Stachl-Peier, Ursula & Tiittula, Liisa 2015. Speech-to-text interpreting in Finland, Sweden and Austria. *Translation & Interpreting* 7(3). 36–49. doi:10.12807/ti.107203.2015.a03.

O'Grady, William. 1987. *Principles of grammar learning*. Chicago: University of Chicago Press.

Pollatsek, Alexander & Hyönä, Jukka. 2005. The role of morphological constituents in reading Finnish compound words. *Language and Cognitive Processes* 20. 261–290.

Potter, Mary C. 1984. Rapid serial visual presentation (RSVP): A method for studying language processing. In Kieras, David Edward & Just, Marcel Adam (eds.), *New methods in reading comprehension research*, 91–118. Hillsdale, N.J: Erlbaum.

Rayner, Keith. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin* 124(3). 372–422.

Rayner, Keith & Duffy, Susan A. 1986. Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory and Cognition* 14. 191–201.

—— 1988. On-line comprehension processes and eye movements in reading. In Daneman, Michelle & MacKinnon, George E. & Waller, Thomas G. (eds.), *Reading research: Advances in theory and practice*, vol. 6, 13–66. New York: Academic Press.

Rayner, Keith & Liversedge, Simon P. & White, Sarah J. 2006. Eye movements when reading disappearing text: The importance of the word to the right of fixation. *Vision Research* 46. 310–323.

Rayner, Keith & Pollatsek, Alexander. 2013. Basic processes in reading. In Reisberg, Daniel (ed.), *The Oxford handbook of cognitive psychology*, 440–461. Oxford: Oxford University Press.

Romero-Fresco, Pablo. 2010. Standing on quicksand: viewers' comprehension and reading patterns of respoken subtitles for the news. In Díaz Cintas, Jorge & Matamala, Anna & Neves, Josélia (eds.), *New insights into audiovisual translation and media accessibility. Media for All* 2, 175–195. Amsterdam: Rodopi.

Räihä, Kari-Jouko & Ovaska, Saila & Sharmin, Selina. 2011. Dynamic text and eye movements: Challenges of tracking and analysis. *Journal of Eye Movement Research* 4(3). 129. doi:10.1016/j.ijhcs.2015.01.010.

Sanders, Deborah S. & Stern, John A. 1980. Effects of text variables on eye movements in reading. *Integrative Psychological and Behavioral Science* 15. 30–34.

Sharmin, Selina & Špakov, Oleg & Räihä, Kari-Jouko. 2012. The effect of different text presentation formats on eye movement metrics in reading. *Journal of Eye Movement Research* 5(3). 1–9.

Sharmin, Selina & Špakov, Oleg, & Räihä, Kari-Jouko. 2015. Dynamic text presentation in print interpreting – An eye movement study of reading behaviour. *International Journal of Human-Computer Studies* 78. 17–30. doi:10.1016/j.ijhcs.2015.01.010.

Sharmin, Selina & Špakov, Oleg & Räihä, Kari-Jouko & Jakobsen, Arnt Lykke. 2008. Where on the screen do translation students look while translating, and for how long? In Gopferich, Susanne & Jacobsen, Arnt Lykke & Mees, Inger M. (eds.), *Looking at Eyes: Eye-Tracking Studies of Reading and Translation Processing* (Copenhagen Studies in Language 36), 31–51. Copenhagen: Samfundslitteratur.

Sharmin, Selina & Wiklund, Mari. 2014. Gaze behaviour and linguistic processing of dynamic text in print interpreting. In *Proceedings of the symposium on Eye Tracking Research and Applications (ETRA'14)*, 263–266. New York: ACM. doi:10.1145/2578153.2578212.

Simola, Jaana. 2011. *Investigating online reading with eye tracking and EEG: The influence of text format, reading task and parafoveal stimuli on reading processes*. Helsinki: University of Helsinki. (Doctoral dissertation.) http://urn.fi/URN:ISBN:978-952-10-7358-8 (Accessed 2015-01-12.)

Špakov, Oleg. 2011. *Sprintanium: A tool for print interpreting*. (Software.) http://www.sis.uta.fi/~csolsp/speechtext/links.php (Accessed 2015-01-13.)

Specker, Elizabeth. 2008. *L1/L2 Eye movement reading of closed captioning: A multimodal analysis of multimodal use.* Tucson: University of Arizona. (Doctoral dissertation.) Available at: http://arizona.openrepository.com/arizona/handle/10150/194820 (Accessed 2016-10-05.)

Staub, Adrian & Rayner, Keith. 2007. Eye movements and on-line comprehension processes. In Gaskell, M. Gareth (ed.), *The Oxford handbook of psycholinguistics*, 327–342. Oxford: Oxford University Press. (http://people.umass.edu/astaub/StaubRayner2007_proof.pdf) (Accessed 2015-01-12.)

Tiittula, Liisa. 2006. Schriftdolmetschen − Mündlichkeit im schriftlichen Gewand [Speech-to-text interpreting – Orality in written form]. In Suominen, Mickael & Arppe, Antti & Airola, Anu & Heinämäki, Orvokki & Miestamo, Matti & Määttä, Urho & Niemi, Jussi & Pitkänen, Kari K. & Sinnemäki, Kaius (eds.), *A Man of Measure: Festschrift in Honour of Fred Karlsson on his 60th Birthday. A special supplement to SKY Journal of Linguistics* 19, 481–488. (http://www.linguistics.fi/julkaisut/SKY2006_1/1FK60.1.10.TIITTULA.pdf) (Accessed 2016-10-15.)

—— 2009. SpeechText: Research on print interpreting. In *Proceedings of 2nd International Seminar on Real-Time Intralingual Subtitling*, Universitat

Autònoma de Barcelona, June 19, 2009. (http://www.respeaking.net/programme/Tiittula.pdf) (Accessed 2013-10-08).
van Dijk, Teun A. & Kintsch, Walter. 1983. *Strategies of Discourse Comprehension*. New York: Academic Press.
Vitu, F. & O'Regan, J.K. & Mittau, M. 1990. Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics* 47(6). 583–600.
Weger, Ulrich W. & Inhoff, Albrecht W. 2006. Attention and eye movements in reading inhibition of return predicts the size of regressive saccades. *Psychological Science* 17(3). 187–191.

Contact information:

Selina Sharmin
Mannikonkatu 4 A 2
33820 Tampere
Finland
e-mail: selina(dot)sharmin(at)gmail(dot)com

Mari Wiklund
Helsinki Collegium for Advanced Studies
Fabianinkatu 24
P.O. Box 4
00014 University of Helsinki
Finland
e-mail: mari(dot)wiklund(at)helsinki(dot)fi

Liisa Tiittula
Department of Modern Languages
P.O. Box 24
000140 University of Helsinki
Finland
e-mail: liisa(dot)tiittula(at)helsinki(dot)fi