Gaze Behaviour and Linguistic Processing of Dynamic Text in Print Interpreting

Selina Sharmin
School of Information Sciences
University of Tampere, Finland
{firstname.lastname}@uta.fi

Mari Wiklund
Department of Modern Languages
University of Helsinki, Finland
{firstname.lastname}@helsinki.fi

Abstract

Print interpreting is a form of communication that allows deaf and hard of hearing people to get access to speech. We carried out an eye tracking experiment where twenty participants read print interpreted text presented dynamically on a computer screen. We compared regression landing points on reread words between two dynamic text presentation formats: letter-by-letter and word-by-word. Then we investigated the gaze behaviour from a linguistic point of view in order to discover whether the dynamic presentation has an effect on linguistic factors. In particular, we have examined the parts of speech of the first and the second landing points of regressions. The findings suggest significant difference between the presentation formats. There is also a relationship between the gaze behaviour and the linguistic processing of dynamic text. Being conscious of this lexical hierarchy may help to develop supporting print interpreting tools and consequently may also help print interpreters to improve the presentation of dynamic text to the user.

CR Categories: H.5.2 [Information Interfaces and Presentation]: User Interfaces – Evaluation/methodology; Input devices and strategies

Keywords: regression, gaze behaviour, linguistic processing, print interpreting, eye tracking, eye movement, dynamic text

1 Introduction

Gaze behaviour has been used in many studies to investigate human cognitive processes. Analysis of eye movements along with linguistic features of the text can be used to recognize indirectly the current mental state of the reader.

Dynamic text presentation formats are often used in various large and small screen electronic devises. Text proceeds automatically in dynamic text presentation format. Bernard et al. (2001) compared three text presentation methods: word-by-word, three-line and 10-line text formats. Reading comprehension was better with the word-by-word and 10-line formats. Studies by Chien and Chen (2007) and Lin and Shieh (2006) have shown that the reading comprehension is better with the word-by-word format than with a letter-by-letter format in reading Chinese text. None of the previous studies have compared dynamic text presentation techniques for print interpreting. Print interpreting is a form of communication that allows deaf and hard of hearing people to get access to speech (Tiittula 2009). In print interpreting, the interpreter translates the speech including significant audible information into written format in real-time. In this study, we compare the gaze behaviour in reading print interpreted dynamic text between two presentation formats, word-by-word and letter-by-letter.

Moreover, the current study investigates the gaze behaviour from a linguistic point of view. The paper focuses on the regressions of eye movements and on their relationship to linguistic factors. The aim of this study is to discover if regressive eye movements vary between the mentioned two dynamic text presentation formats, and if there is any relationship to linguistic factors. The analyses are theoretically based mainly on O'Grady's (1987) cognitive approach to the acquisition and to the use of language.

In O'Grady's (1987) cognitive approach, nouns are "primary" elements of language, because they are characterized by autonomous meaning and function. Their referents are perceptively distinct and coherent. 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, Furtner et al. 2009). Verbs and adjectives are "secondary" elements. They depend on a relationship to at least one primary element. Function words, in turn, are "tertiary" elements. They depend on a relationship to at least one secondary element. Because of this lexical hierarchy, children tend to learn first most of all nouns, then verbs and adjectives, and the function words are the last ones to be acquired (Furtner et al. 2009).

Gaze behaviour in reading has been analysed in several previous studies (Rayner 1998). Typical fixation duration in reading is 100–500 ms. However, readers do not fixate all the words. Foveal processing of each word is not necessary. Especially many short words are skipped (Weger and Inhoff 2006). Previous studies have shown that content words are fixated about 85% of the time, whereas function words are fixated only about 35% of the time (Carpenter and Just 1983; Rayner and Duffy 1988, Furtner et al. 2009).

In reading, eyes do not move forward persistently all the time; they also move backwards for rereading. Opposite movements of eyes from right-to-left along the line or movements back to previously read words and lines are called regressions.

The first landing point is the first word where the regression lands at. The second landing point is the second word where the regression lands at. Short within-word regressions may occur when the reader has difficulty in processing the currently fixated word (Carpenter and Just 1983; Rayner and Duffy 1988). Longer regressions back along the line or to another line may occur because the reader did not understand the text (Rayner 1998).

In this study, we have examined the parts of speech of the first and the second landing points of regressions. Then we have classified them into primary, secondary and tertiary elements. The objective is to discover if the same lexical hierarchy that prevails in the acquisition of language can also be found here.

We examined eye tracking data from an experiment where twenty test participants read a dynamic text presentation on a computer screen. In our data, the text is a spoken presentation that has been transformed into written format. This may play a role in the gaze behaviour. Results of this study reveal that gaze behaviour differs due to different presentation formats. More rereading occurs in the word-by-word format compared to the letter-by-letter format although the number of reread words per regression was almost equal for both presentation formats. From the linguistic point of view, the findings suggest that there is a relationship between the gaze behaviour and the linguistic processing of dynamic text. Understanding this lexical hierarchy may help to develop supporting print interpreting tools and consequently may also help print interpreters to improve the presentation of dynamic texts to the user.

2 Background

Eye-tracking is a sensitive method to language processing without interfering with it. It allows non-disruptive observations in natural experimental settings. Previously many studies have tested the participant's ability to maintain information from the perspective of reading span tasks. It has been shown that elaborative encoding strategies – such as chaining, mental imagery and semantic elaboration – are more beneficial than simple rehearsal (Daneman and Carpenter 1980; Friedman and Miyake 2004; Kaakinen and Hyönä 2007; Turner and Engle 1989; McNamara and Scott 2001). The development of general theories of language processing also made it possible to use eye movements for examining cognitive processes underlying reading (Rayner 1998). Tanenhaus and Trueswell (2006) gave an overview of research that uses eye movements to investigate spoken language comprehension. The study was about spoken language, but nothing was discovered about transcribing spoken language into written format.

Different factors, such as a word's frequency, length, predictability, and ease of integration into the sentence are believed to influence eye movements on the particular part of the text. Those factors influence whether and for how long the eyes fixate on a word (Just and Carpenter 1980; Rayner 1998; Reichle et al. 2003). Readers develop structural (syntactic) representation of sentences incrementally in reading each word. Usually detection of syntactic or semantic irregularities in a word evokes longer fixations, regressive eye movements, and rereading (Altmann et al. 1992; Ferreira and Henderson 1990).

Rereading is a natural human behaviour of eye movements in reading. It can indicate an active process to serve 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 go well, 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 and Nurminen 2006). Thus, different eye movement behaviours in reading, such as looking at the text for a long time or creating longer fixations or looking back, could be caused by different cognitive mechanisms.

3 Methods

3.1 Participants

Twenty native Finnish speaking participants took part in the experiment. All participants had normal or corrected-to-normal vision. The average age of the participants was 28.4 years, with a standard deviation (SD) of 8.9 and an age range of 18–51 years. All of them were either members of the university staff or students.

3.2 Apparatus

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

3.3 Procedure and Design

First, participants were informed about the test procedure. Then the eye tracker was calibrated for the participants' eyes. The distance between the monitor and the participant was about 60 cm. Stimulus consisted of a text (160 words in length), which was the output of a print interpreting process, where a professional print interpreter was transforming a spoken conference-like presentation into written format in Finnish. The interpretation was first produced in a live situation, and afterwards rendered at real-time pace on the screen either in the word-by-word presentation format or in the letter-by-letter format. Half of the participants read the text presented in word-by-word format while the other half read it in the letter-by-letter presentation format.

A background questionnaire was delivered at the beginning of the test. There was also a post-test questionnaire regarding user experience about reading the text. Participants were informed about the post-test questionnaire in order to motivate them to read the text carefully.

4 Results

Videos of the eye movements of each participant were used in the analysis. A careful observational analysis was carried out to spot the words from which the rereading or the regressions started as well as the words where the regressions landed at.

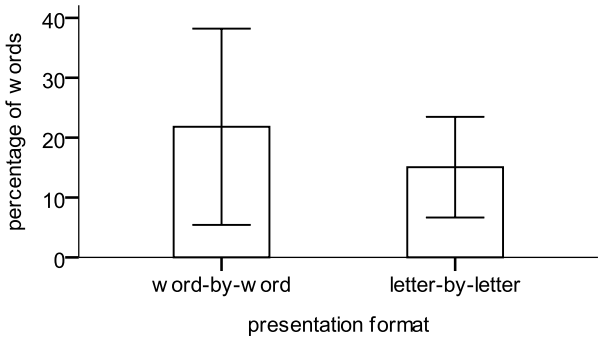

Figure 1: Percentage of words from which rereading started. The error bars denote the standard deviation.

Percentages of words from which rereading started for the word-by-word and the letter-by-letter presentation formats were 21.81

and 15.06, respectively. One way ANOVA showed that the difference was significant with $p < .05$, $F_{1,19} = 5.373$ (Figure 1). Thus, participants started to reread more in the word-by-word presentation format. Moreover, the percentage of regression landing points in the word-by-word presentation format (83.25) was significantly more than in the letter-by-letter format (50.63) with $p = .05$, $F_{1,19} = 4.434$. Thus, participants were more likely to start rereading and consequently there were more regression landing points in the word-by-word presentation format than in the letter-by-letter presentation format. On the other hand, it is interesting that although there were more regression landing points in the word-by-word format, regression length or percentage of average number of reread words per regression was almost equal for both presentation formats (average values for the letter-by-letter and the word-by-word formats were 3.44 and 3.73, respectively).

We continued our analysis from the linguistic point of view by observing the landing points per regression in the letter-by-letter presentation format. We examined the words at which the first two regression points landed. Our data consists of 109 regression clusters. As all the first landing points are followed by second landing points, the total number of the first and the second landing points is 218 (= 109 + 109). We have examined the parts of speech of these first and second landing points of regressions. Then we have classified them into primary, secondary and tertiary elements. Our linguistic analysis is theoretically based mainly on O'Grady's (1987) cognitive approach to the acquisition and to the use of language.

The results show that the first landing point is a noun (that is, a primary element) in about 52% of the cases. For instance in Example 1, the first landing point is the word 'ryhmittymiä' ('groups'), which is a noun. The text in Example 1 is presented in Figure 2 with landing points as clusters of circles. Numbers inside the circles indicate the ascending sequence of gaze points.

Example 1:
'In America there are even religious **groups** that champion this **cause**.'
→ The first landing point: *ryhmittymiä* (noun)

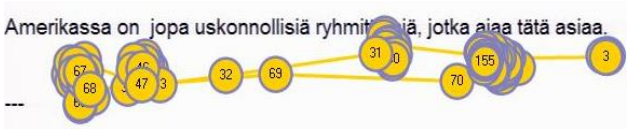

Figure 2: Example of a noun as the first landing point

When we studied the landing points in more detail, we discovered that the first landing point is the closest noun in 33% of the cases.

Example 2:
'It refers to developmental disorders of the central nervous system, including autism and **aspegren syndrome** *[sic]*.'
→ The first landing point: *aspegren* (the closest noun)

The first landing point is a content word (that is, a primary or a secondary element) in 76% of the cases.

Example 3:
'There has been a lot of discussion about MPM *[sic]* -vaccines (?) that they **would cause autism**.'
→ The first landing point: *tuottais*, 'would cause' (a verb → a content word)

In Example 3, the first landing point is the verb tuottais ('would cause'), which is a content word. The first landing point is also the closest content word here. This is the case in 43% of the regressions.

When we considered the first and the second landing points of the regressions, we obtained the following results. The first or the second landing point of the regression is a noun in 74% of the cases, even if less than 30% of all words of the data are nouns. This is in line with the findings of Furtner et al. (2009) according to which readers recur their fixations to nouns more than to words of other parts of speech in order to enhance the comprehension of the surrounding words. When the first or the second landing point is not a noun, it is most likely a verb or an adjective: indeed, the first or the second landing point is a content word in 91% of the cases. This is interesting, because only 57% of all words of the data are content words. Function words, in turn, are rare in this position: the first or the second landing point is a function word only in 9% of the cases, even if as much as 43% of all words of the data are function words.

5 Discussion and Conclusions

Earlier studies have shown that the comprehension is improved in the word-by-word format compared to the letter-by-letter format. On the other hand, Rayner (1998) has documented that look backs or rereading can be indicators of comprehension difficulties. In this study, observing gaze behaviour revealed significantly more regressive starting points and landing points, i.e., participants reread previous words more, in the word-by-word format than in the letter-by-letter format. The increase in number of regressions can have two causes: poor comprehension on first reading, or desire to strengthen comprehension when there is a better chance for it because the new text appears with longer breaks between words than between letters. In either case regressions should help and result in better understanding. In addition, regression length or percentage of reread words per regression was almost equal for both presentation formats. Hence, in parallel to existing default letter-by-letter presentation in print interpreting, this study suggests the use of the word-by-word format as an alternative.

As already mentioned, previous studies have shown that content words are fixated about 85% of the time, whereas function words are fixated only about 35% of the time (Carpenter and Just 1983; Rayner and Duffy 1988, Furtner et al. 2009). This result falls in line with our findings according to which the first and the second landing points of regressions are generally (90.8%) content words.

Indeed, our results show that the test subjects look for primary and secondary elements of language in order to construct the meaning of what they have just read. Nouns, which are primary elements, are the most likely landing points of regressions (Furtner et al. 2009).

The fact that the lexical hierarchy that can be found here is the same as the one typically observed in the acquisition of language (O'Grady 1987; Furtner et al. 2009) reflects the cognitive processing of language by which the meaning is being constructed. This, in turn, suggests that there is a relationship between the gaze behaviour and the linguistic processing of dynamic text. In addition, since regressions were more common for the word-by-word presentation format, rendering the text word-by-word should be supported by print interpretation software. Sprintanium (Špakov 2011), the tool used in our

experiments, does this, whereas many professional print interpreters use simply Microsoft Word which does not have this option.

Being conscious of this lexical hierarchy may also help print interpreters to improve the presentation of dynamic texts to the user. This could be done for example by highlighting the primary elements and by reducing only tertiary elements of language.

References

ALTMANN, G. T. M., GARNHAM, A., AND DENNIS, Y. 1992. Avoiding the garden path: Eye movements in context. *Journal of Memory and Language 31*, 685–712.

BERNARD, M. L., CHAPARRO, B. S., AND RUSSELL, M. 2001. Examining automatic text presentation for small screens. In *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, Minneapolis, 637–639.

CARPENTER, P. A., AND JUST, M. A. 1983. What your eyes do while your mind is reading. In Rayner, K. (ed.), *Eye movements in reading: Perceptual and language processes*. Academic Press, 275–307.

CASELLI, M., BATES, E., CADADIO, P., FENSON, J., SANDERL, L., AND WEIR, J. 1995. A crosslinguistic study of early lexical development. *Cognitive Development 10*, 159–199.

CHIEN, Y.-H, AND CHEN, C.-H. 2007. The use of dynamic display to improve reading comprehension for the small screen of a wrist watch. In *Proceedings of the 2007 Conference on Human Interface, Part I*. Springer, 814–823.

DANEMAN, M., AND CARPENTER, P. A. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour 19*, 450–466.

FERREIRA, F., AND HENDERSON, J. M. 1990. The use of verb information in syntactic parsing: A comparison of evidence from eye movements and self-paced word-by-word reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition 16*, 555–568.

FRIEDMAN, N. P., AND MIYAKE, A. 2004. The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language 51*, 136–158.

FURTNER, M. R., RAUTHMANN, J. F., AND SACHSE, P. 2009. Nomen est omen: Investigating the dominance of nouns in word comprehension with eye movement analyses. *Advances in Cognitive Psychology 5*, 91–104.

GENTNER, D. 1982. Why nouns are learnt before verbs: linguistic relativity versus natural partitioning. In Kuczaj, S. (ed.), *Language development, vol. 2*. Erlbaum, 67–88.

HYÖNÄ, J., AND NURMINEN, A. M. 2006. Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology 97*, 31–50.

HYÖNÄ, J., LORCH, R. F., AND KAAKINEN, J. 2002. Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology 94*, 44–55.

JUST, M. A., AND CARPENTER, P. A. 1980. A theory of reading: From eye fixations to comprehension. *Psychological Review 87*, 4, 329–354.

KAAKINEN, J., AND HYÖNÄ, J. 2007. Strategy use in the reading span test: An analysis of eye movements and reported encoding strategies. *Memory 15*, 6, 634–646.

LEVY, B. A., DIPERSIO, R., AND HOLLINGSHEAD, A. 1992. Fluent rereading: Repetition, automaticity, and discrepancy. *Journal of Experimental Psychology: Learning, Memory and Cognition 18*, 957–971.

LIN, Y.-C., AND SHIEH, K.-K. 2006. Reading a dynamic presentation of Chinese text on a single-line display. *Displays 27*, 145–152.

MARATSOS, M. 1991. How the acquisition of nouns may be different of that of verbs. In Krasnegor, N., Rumbaugh, D., Schiefelbusch, R. and Studdert-Kennedy, M. (eds.), *Biological and behavioral determinants of language development*. Erlbaum.

MCNAMARA, D. S., AND SCOTT, J. L. 2001. Working memory capacity and strategy use. *Memory and Cognition 29*,10–17.

O'GRADY, W. 1987. Principles of grammar learning. University of Chicago Press.

RAYNER, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin 124*, 372–422.

RAYNER, K., AND DUFFY, S. A. 1988. On-line comprehension processes and eye movements in reading. In Daneman, M., MacKinnon, G. E. and Waller, T. G. (eds.), *Reading research: Advances in theory and practice*. Academic Press, 13–66.

REICHLE, E. D., RAYNER, K., AND , A. 2003. The E-Z Reader model of eye movement control in reading: Comparison to other models. *Brain and Behavioral Sciences 26*, 445–476.

ŠPAKOV, O. 2011. Sprintanium: A tool for print interpreting. Software URL: http://www.cs.uta.fi/speechtext/links.php.

TANENHAUS, M. K., AND TRUESWELL, J. C. 2006. Eye movements and spoken language comprehension. In Traxler, M., and Gernsbacher, M. (eds.), *Handbook of Psycholinguistics (2nd ed.)*. Elsevier, 863–900.

TIITTULA, L. 2009. SpeechText: research on print interpreting. In *Proceedings of 2nd International Seminar on Real-Time Intralingual Subtitling*, Universitat Autonoma de Barcelona. URL: www.cs.uta.fi/speechtext/docs/Tiittula_ISRIS_2009.pdf

TURNER, M. L., AND ENGLE, R. W. 1989. Is working memory task dependent? *Journal of Memory and Language 28*, 127–154.

WEGER, U., AND INHOFF, A. 2006. Attention and eye movements in reading inhibition of return predicts the size of regressive saccades. *Psychological Science 17*, 3, 187–191.