



MultiplEYE

The 2nd Workshop on Eye Movements and
the Assessment of Reading Comprehension

Zurich, 20-22 June

Book of Abstracts

Editors:

Yevgeni Berzak (Technion – Israel Institute of Technology)

Titus von der Malsburg (University of Stuttgart)

Lena A. Jäger (University of Zurich)



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Keynote

Author: Jean-François Rouet

Title: Reading as a Purposeful Activity: Some Implications for the Online Assessment of Comprehension

Abstract:

Reading involves a set of skills that enable competent readers to engage with any type of text for any type of purpose. Typical reading purposes include comprehending a text passage as a whole, but also locating information of interest within the passage, using text information to reason or to make decisions, and even comparing and contrasting information from multiple texts. These purposes each call for specific reading strategies. The RESOLV theory of reading (Britt et al., 2018; Rouet et al., 2018) considers readers' production of task goals, and their ability to make goal-relevant decisions during reading as a key dimension of comprehension proficiency. Such a purpose-driven approach to reading provides new insights for the study and online assessment of reading. I review some examples based on research studies conducted with children, teenagers and young adults. Then I suggest some perspectives for future research.

References:

Britt, M. A., Rouet, J. F., & Durik, A. (2017). Literacy beyond text comprehension: A theory of purposeful reading. Routledge.

Rouet, J. F., Britt, M. A., & Durik, A. M. (2017). RESOLV: Readers' representation of reading contexts and tasks. *Educational Psychologist*, 52(3), 200-215.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Talks Day 1

Author(s): Maja Stegenwallner-Schütz, Deborah Jakobi, Ramuné Kasperè, Nora Hollenstein

Title: Developing a multilingual eye-tracking-while-reading corpus

Abstract:

The MultipleYE reading corpus is a multilingual eye-tracking-while-reading corpus with healthy, adult, L1 readers collected in a large-scale international multi-lab effort. At the moment, the project includes more than 20 languages of which many are understudied. The project is based on four core principles which are consistency, open access, diversity, and transparent data quality reporting. Following these principles, the experiment design and data pre-processing pipeline is consistent across languages and collection sites to enable optimal cross-language comparison and flexible use of the data. Furthermore, we are establishing an entirely open source pipeline for eye-tracking data that ensures transparency and reproducibility, and allows for merging data sets recorded with different hardware. All data, software, protocols, and metadata will be made available via the MultipleYESTore (hosted by PsychArchives). To enable maximal diversity there are only little requirements to contribute. For example, only little restrictions are set on eye-tracking devices, or there is the possibility to contribute additional datasets testing different populations or using different stimuli or methods. In order to ensure transparent data quality reporting, one aim of the project is to develop new standards for reporting data quality and how to document.

A core pillar of the data collection is to design reading comprehension questions. The questions target different levels of cognitive processes that are required to fully understand the presented text. This will allow for fine-grained analyses of the readers' understanding of the text. The questions undergo a thorough and iterative reviewing process to avoid common shortcomings of comprehension questions such as ceiling effects.

The vision of this newly created corpus is not only to allow for psycholinguistic analyses of cognitive processes underlying reading but also to allow for a wide variety of other research goals such as exploring how reading comprehension can be assessed using state of the art machine learning models or studying human and large language model reading comprehension performance.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Diane C. Mézière, Lili Yu, Titus von der Malsburg, Erik D. Reichle, Genevieve McArthur

Title: Predicting Recall from Eye Movements during a Read-Only Task

Abstract:

Research on the usefulness of eye-movement measures as predictors of reading comprehension has shown that while such measures can indeed predict comprehension scores on various reading comprehension tasks (e.g., questions), the relationship between eye-movement measures and comprehension scores is mediated by differences in task demands across comprehension measures. One possible advantage of eye-tracking is that it does not require an overt additional comprehension task such as questions. Hence, it may be possible to use eye movements during a reading task with no such additional demands to predict comprehension ability. In this study, we investigate the usefulness of eye movements collected during a reading task with no additional demands to predict reading comprehension as measured by recall. We tested 62 adult native English speakers. Participants were given 9 passages to read while we tracked their eye movements. These 9 passages were divided into 3 sets, according to three experimental conditions in which the task was manipulated: read-only, recall-1, and recall-2. In the read-only condition, participants were asked to read the text for comprehension. In the recall-1 and recall-2 conditions, participants were told to read the passages and then asked to recall the contents as accurately as possible. Passages from the read-only and recall-1 conditions were counter-balanced. The recall task was scored using propositional text bases. The relationship between eye-movement measures and recall was examined using Bayesian linear models and cross-validation to identify the best predictors of recall performance. Specifically, we predicted recall-2 scores with eye movements from a) the read-only condition and b) the recall-1 condition. The outputs of the top 10 models for each analysis were then compared to examine the usefulness of eye-movements collected in the read-only condition to predict recall. Results showed that both sets of eye movements were equally useful in predicting recall performance ($R^2 = 0.26$ and 0.27 respectively). Additionally, the most useful predictors of recall performance were nearly identical across conditions. These results demonstrate the usefulness of eye-movement measures collected during a reading task with no additional demands to predict reading comprehension.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Kars Ligtenberg, Stefan Frank, Sybrine Bultena

Title: Using Eye-Tracking to Quantify L2 Proficiency in Dutch-English Bilingual Sentence Processing

Abstract:

The language systems of multilinguals are not independent or separate systems that can be turned on or off at will, instead they form an integrated system. How this system operates, and what mechanisms underly its operation, are still largely unexplored. We propose that statistical regularities of language guide many language-processing phenomena, including syntactic transfer between the languages available to a bilingual. With our experiment we aim to uncover when and how statistical properties of language transfer between languages by co-registering eye-tracking and EEG. Co-registration of these two paradigms allows for natural reading, otherwise uncommon in EEG research, and has previously proven to be able to reveal effects of language statistics during reading.

We aim to do this with the language-pairing of Dutch and English, as these languages are closely related, which should facilitate transfer, and yet have some specific grammatical differences that can be clearly denoted, which ensures transfer effects and directions thereof can be pinpointed. Importantly, we will test both L1 Dutch and L1 English participants, and vary the levels of exposure of Dutch-English bilinguals through recruiting from different student populations (i.e., those enrolled either in English-language or in Dutch-language programmes).

However, exposure alone does not guide potential differences in reading measures, so we aim to use the eye-tracking data collected during the co-registered experiment to also ascertain level of L2 proficiency. Berzak, Katz & Levy (2018) has shown that eye-tracking measures can reliably indicate reading proficiency and even seemed to correspond rather well with MET (Michigan English Test) and TOEFL (Test Of English as a Foreign Language) test scores. Additionally, one of the measures used by Berzak et al (2018) was Word Property Coefficients, which aside from eye-tracking measures, also relied on surprisal. Our project uses surprisal, and relative differences in surprisal, as one of the outcome measures of language transfer effects: this might be able to further enhance the use of eye-tracking (or even eye-tracking co-registered with EEG) as a measure of reading proficiency.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Grant Eckstein, Troy Cox

Title: Eye Movements as Proficiency Indicators in Reading Across Four Languages

Abstract:

Measuring language proficiency in a second or foreign language through reading is a difficult proposition. Typical reading assessments evaluate student proficiency through multiple choice testing, short answer, summary/recall, or even reading speed tasks. Yet these types of assessments are time-consuming and effortful to create and administer, and they only measure the output of reading rather than the process of reading directly. Yet eye-tracking has the potential to directly measure the behaviors of reading that index reading processes of more and less proficient readers in a language.

We are interested in using eye-tracking measures to develop a proficiency-based reading model that can then be used for assessment purposes. This is possible because the speed at which readers access lexical knowledge can be correlated with certain eye-tracking measures. At a very broad scale, this is illustrated by readers making fewer fixations, shorter saccades, fewer regressions, and more skips as their language proficiency increases. Extremely fine-grained eye-movement data can better model proficiency-based reading behavior. Moreover, research has demonstrated that factors such as language typology and word frequency in the lexicon modulate how quickly readers can access lexical information, so controlling for these and other variables can further increase the accuracy of a proficiency-based reading model.

In this presentation, we will describe the efforts of our transdisciplinary group to collect eye movement data and build a large language model that can assess a reader's language proficiency based on eye movements during naturalistic reading. So far, we have collected reading behaviors of over 100 L2 learners with various levels of reading proficiency in four languages (English, Chinese, Russian, and Portuguese). Each learner read passages ranging from Intermediate to Superior in level. We explored numerous eye-tracking measures as participants read below, at, and above their reading proficiency. Applying machine learning techniques to Russian pilot data revealed that readers can be classified into broad proficiency categories (beginner vs. advanced) with 87% accuracy based on their eye movements when reading a single text. When considering all texts, accuracy approached 100%. These early results are promising and suggest that proficiency assessment via eye tracking is indeed feasible.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Erik Reichle, Signy Wegener, Elisabeth Beyersmann, Hua-Chen Wang, Kate Nation, Anne Castles

Title: Eye Tracking and Simulating the Spacing Effect during Orthographic Learning

Abstract:

Purpose: Spreading out study opportunities over time improves the retention of verbal material compared to consecutive study, yet little is known about the influence of temporal spacing on orthographic learning specifically. The current study addressed four questions: (1) do participants' eye movements during orthographic learning differ under spaced and massed conditions?; (2) is the spacing effect observed in off-line post-tests?; (3) can participants eye movements during learning be linked with their learning success in off-line post-tests?; (4) can E-Z Reader (Veldre et al., 2023) simulate the spacing effect during orthographic learning?

Method: Eighty adults silently read contextual sentences containing novel words while their eye movements were monitored. Sentences were read four times; half of the items were spaced while half were massed. Participants completed a post-test assessing their written word form learning (orthographic choice or spelling-to-dictation). Simulations with E-Z Reader were used to interpret the human data.

Results: During orthographic learning, massed items had shorter total reading times than spaced items. A clear spacing advantage was noted in the offline post-tests. Longer fixations during learning were associated with higher response accuracy at post-test. Implementing a processing deadline enabled E-Z Reader to simulate participants' eye movements as they encountered novel words during reading; simulations of the spaced and massed conditions suggested that the latter items may have received less attentional processing than the former items during learning.

Conclusions: Spreading out experiences with new written words results in longer fixations during learning and better learning outcomes using off-line tests. The combination of human eye movements and computational modelling provides useful insights into how reading and memory intersect and point to new directions for future research.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Talks Day 2

Author(s): Carolin Hahnel, Robin Merchel, Lothar Persic-Beck, Frank Goldhammer

Title: Using Eye-Tracking to Trace Processes of Information Integration and Sourcing in Multiple Document Reading

Abstract:

Learners often use information that is spread across multiple digital texts or information sources, especially when learning with Internet texts. Successful learning with multiple documents requires learners to identify relevant information, interpret it appropriately according to its origin or intent, and compare and integrate it with other information from alternative sources. In our study, we focus on the two central processes of multiple document reading: the integration of information across texts and the use of metadata (sourcing). Using eye-tracking, we investigate the extent to which these processes occur when processing multiple documents and how they can be represented through eye-tracking data.

For this purpose, we use four released units of the digital PISA reading competence assessment. As a baseline, all participants first work on one unit (7 reading comprehension items relating to a forum exchange on chicken health). Subsequently, the participants randomly receive one of two short interventions in which they receive prompts targeted at one of the two focused processes of multiple document reading: One group receives a worked example with a focus on intertextual comparison of information, while the other group receives a worked example with a focus on processing meta-information of the texts. Afterwards, the participants are asked to work on three further PISA reading units (Cow's Milk, Rapa Nui, The Galapagos Islands; 3 items each). We expect that learners in the first group will compare text passages across documents more intensively than learners in the second group, who are asked to consider the meta-information in the texts more intensively. Accordingly, we expect that after the prompts have been studied, learners in the first group will switch more frequently between relevant text passages in the stimuli, while learners in the second group should pay more attention to meta-information. Reading speed is also recorded as an important covariate.

We are currently recruiting a sample of school students in grades 8-10 as well as a sample of university students. We will present the first results of the study at the workshop.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Lili Yu, Erik Reichle

Title: The Effects of Auditory Information and its Rate on Reading

Abstract:

Reading research has been primarily focused on the reading of static text in the absence of other distractions. However, in real life, reading often takes place in the presence of background speech, such as trying to read text on a presentation slide while the speaker is simultaneously giving the talk. In this study, we investigated how the simultaneous presentation of auditory speech along with text affects online reading and its comprehension. In the study, the presence vs. absence of identical auditory speech and its rate were manipulated, along with the difficulty of sentences being read. Sixty-two university students' eye movements were monitored while they read 192 sentences in four experimental conditions: (1) text only; and text with identical auditory speech presented at a rate that was (2) faster than; (3) equivalent to; or (4) slower than participants' self-paced reading speed. Two types of sentences (easy vs. difficult, with difficult sentences containing garden-path structures) were included to examine if the text-auditory integration is conditioned upon the sentences being read. Half of the sentences were followed by comprehension questions. Our results showed that participants adapted their eye movements to accommodate the different speech rates. In comparison to silent reading, participants made longer fixations when the speech rate was slow, fewer but longer fixations when the speech rate was similar to participants' reading speed, and fewer and shorter fixations with the fast speech rate. Interestingly, when the auditory speech was presented, participants' reading comprehension improved irrespective of the speech rate (compared to the text-only condition) or sentence types. Our study demonstrates that (even skilled) readers make use of all available information resources to facilitate their reading comprehension; they also adapt their eye movements to accommodate the rate of auditory speech. These results provide important clues about the mental processes that support comprehension in many real-life multi-modal reading situations (e.g., reading subtitles in film).



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Omer Shubi, Yoav Meiri, Yevgeni Berzak

Title: Fine-Grained Prediction of Reading Comprehension from Eye Movements

Abstract:

Can human reading comprehension be assessed directly from eye movements without presenting reading comprehension questions? In this work, we address this longstanding question using a combination of multi-modal language modeling and large-scale eyetracking data over textual materials that are geared towards behavioral analyses of reading comprehension. Differently from prior work that aggregated responses over multiple items, we focus on a challenging and largely unaddressed task of predicting reading comprehension of a single participant with respect to a single question about one passage. We examine this task in different evaluation setups and in two reading conditions: ordinary reading, and information seeking, a common reading scenario in human question-answering.

This task is made possible by OneStop, the largest eye-tracking dataset for reading comprehension in English to date, containing 486 multiple choice reading comprehension questions and 19,440 responses from 360 native English speakers. In the ordinary reading regime participants receive the question only after having read a passage, while in the information seeking regime they are also presented with the question prior to reading the passage. We introduce three multi-modal transformer encoder models based on the BERT architecture that correspond to three different ways of combining text and eye movement data to predict reading comprehension. We further implement several models from the literature as well as a BERT text-only baseline. We evaluate the models in three regimes: predicting comprehension for new items and seen participants, new participants and seen items, and both new items and participants. Furthermore, we leverage a categorization of distractors by degree of comprehension to extend the standard binary evaluation (correct vs. incorrect) to a finer-grained analysis of distractor choices. Preliminary results indicate that the task at hand is highly challenging for the models examined compared to the text-only baseline. The data, tasks, models, and evaluation schemes presented here will serve as benchmarks for future research on NLP-driven eye movement modeling and automatic assessment of reading comprehension.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Yoav Meiri, Yevgeni Berzak

Title: Eye Movements and Reading Comprehension in Repeated Reading

Abstract:

This study investigates eye-movement differences between first and repeated reading of the same text, and how these differences relate to reading comprehension. By analyzing a large eye-tracking dataset of L1 English readers, we examine repeated reading effects under different reading regimes and explore individual differences in reading speed and comprehension levels.

We utilized OneStop, a broad coverage eye-tracking dataset, which includes 360 L1 English participants. Each participant read 10 Guardian articles in one of two reading regimes: ordinary reading or information seeking, and answered a multiple-choice question after each paragraph. For each participant, two of the articles were presented for a second time. We analyzed standard eye movement measures and their sensitivity to word properties in repeated reading compared to first reading. Linear mixed-effects models were used to assess the effects of reading regime, relevance of information to the task, and individual differences.

Repeated reading led to improved reading comprehension in both reading regimes. This was accompanied by shorter reading times, fewer fixations and regressions, and reduced sensitivity to word properties compared to first reading, suggesting more efficient processing. In information seeking, facilitation was greater for task-irrelevant information, and when the division of relevant/irrelevant information was identical across readings, implying efficient encoding of task-specific information. Analysis of individual differences revealed that slower readers exhibited larger facilitation effects in repeated reading. Contrary to this outcome, the repeated reading facilitation did not depend on reading comprehension performance in the first reading, suggesting that repeated reading benefits are driven by the amount of prior text processing, rather than its effectiveness.

These findings provide an empirical characterization of the effects of repeated reading and their relation to reading speed and reading comprehension during first reading, which could inform future models of eye movements in reading. They further indicate that depending on whether the text is read for the first or second time, very different eye movement patterns can underlay similar reading comprehension outcomes. This poses new challenges and opportunities for developing more fine-grained models for predicting reading comprehension from eye movements, which will take into account the effects of prior exposure to the text.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Margarita Ryzhova, Emilia Ellsiepen, Katharina Trinley, Iza Skrjanec, Vera Demberg

Title: The Effects of Linguistic Context on Comprehension of Unknown Words

Abstract:

Words that are unfamiliar to us can elicit processing difficulties. Word familiarity can be modulated by the intrinsic properties of the word like frequency and length (Rayner, 1998, Kliegl et al. 2004). However, the literature shows that the context also affects comprehension (Nieuwland & van Berkum 2006; Lowell & Morris, 2014; Williams & Morris, 2004). For example, scientific or technical texts may contain more specialized vocabulary that is unfamiliar to the general reader, while everyday texts such as newspapers or novels may contain more familiar language. In such common contexts, the reader can be surprised to encounter an unknown word, or attribute it to a typo, while in a more scientific context, the reader might expect to encounter special domain terms that they don't know.

In our study on processing unknown words in German, we manipulate the type of context to explore whether it affects the reader's sensitivity to processing unfamiliar words. We conduct a self-paced reading experiment and ask participants to read texts for comprehension. Each text includes a target word: either a *real word* or a *pseudoword*. The target words were embedded into two types of context: *everyday* and *scientific*, making this study follow a 2x2 design. Everyday stories concern familiar events from daily life (e.g. children playing in a park), while scientific stories take place in less common settings with characters with a specialized profession (e.g. researchers conducting experiments in a laboratory). The scientific stories themselves are not expository texts, but rather narratives describing a less familiar scenario.

We find that in both contexts subjects showed sensitivity to pseudowords, resulting in higher reading times. However, this effect was significantly stronger in the everyday context, compared to the scientific context condition. The context alone didn't affect the reading times. Our results show that unknown words, despite lacking defined meaning, are more anticipated in domain-specific texts than in general narratives. The scientific context increases the expectancy of encountering unknown words, resulting in faster reading.

In the time of abstract submission, we are conducting an eye-tracking counterpart of this study, additionally collecting information on language experience and domain expertise.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Jan Brassler, Chiara Tschirner, Maja Stegenwallner-Schütz, Lena Jäger

Title: Deciphering Literacy among Child Learners: Exploring the Link between Reading Comprehension and Visual Search Behaviour through Eye-Movements

Abstract:

The close association between reading comprehension abilities and visual search indicates that skilled readers tend to display more efficient gaze patterns [1, 2]. This relationship is particularly interesting when studying children with limited reading experience, as their gaze behaviour sheds light on developmental trajectories. Yet, the specific influence of reading experience on visual search behaviour and the gaze patterns indicative of strong reading ability levels, especially reading comprehension levels, among children remain largely underexplored. To this end, we conducted an eye-tracking experiment comprising a visual search task on different types of symbols (letter vs. generic shapes) with children aged six to nine at Swiss schools. We identify which features of children's gaze data are most predictive of their word-level reading comprehension abilities, as assessed by a standardized test, employing Bayesian inference methods. The model predicting word-level reading comprehension most accurately includes as predictive gaze features Average Target Fixation Duration, and Durational Difference of Consecutive Non-Target Re-fixations, which are presumably associated with information extraction efficiency [3]. In addition, the model includes Start Left, which is indicative of reading-like search behaviour, and Average Saccade Amplitude, which indicates systematic search behaviour. Furthermore, we highlight how the type of symbols employed in the visual search tasks influences predictive gaze features of reading comprehension scores. Specifically, when models are limited to the letter condition, the most accurate model differs substantially from the general model, sharing only the Average Target Fixation Duration feature. This suggests the importance of adjusting the type of feature used for inferring reading comprehension from visual search depending on the type of stimuli. In summary, our study underscores the intertwined nature of reading comprehension abilities and visual search strategies, particularly in children with limited reading experience. Our findings emphasize how both the selection of symbols and the level of reading experience significantly influence the observed association between visual gaze patterns and reading comprehension abilities.

References:

- [1] Ferretti, G., Mazzotti, S., & Brizzolara, D. Visual scanning and reading ability in normal and dyslexic children. *Behavioural Neurology*, 19(1-2):87–92, 2008.
- [2] Carroll, J. M., Solity, J., & Shapiro, L. R. Predicting dyslexia using prereading skills: The role of sensorimotor and cognitive abilities. *Journal of Child Psychology and Psychiatry*, 57(6):750–758, 2016.
- [3] Rayner, K. The 35th Sir Frederick Bartlett Lecture: Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, 62(8):1457–1506, 2009.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Ekta Sood, Prajit Dhar

Title: NMT-Eye: Towards Multilingual Comprehension with Hybrid Text Saliency Models

Abstract:

Eye movements during reading offer a window into cognitive processes and language comprehension, but the scarcity of reading data with interruptions -- which learners frequently encounter in their everyday learning environments -- hampers advances in the development of intelligent learning technologies. To this end, we recently introduced InteRead (Zermiani et al. (2024)) -- a novel 50-participant dataset of gaze data recorded during self-paced reading of real-world text. InteRead further offers fine-grained annotations of interruptions interspersed throughout the text as well as resumption lags incurred by these interruptions. Our analyses show that our dataset is in line with relevant previous findings (Schilling et al. (1998), Kiegl et al. (2004)), particularly regarding first pass fixations. Besides, we find that the logarithmic token frequency exhibits a non-linear relationship with word length, as found in Kiegl et al. (2004).

In this Workshop, we plan on presenting our findings as well as starting a discussion on how to handle zero-count fixation data. Fixations, measured either as durations or raw counts, follow an extremely skewed distribution. In our experiments, we found 59% of the fixation counts to hold a value of zero, with the distribution being bimodal. While the standard approach in dealing with zero-count fixations entails setting a lower-bound, for instance discarding fixation durations less than 100ms, we propose to retain the fixation data as it is, but instead utilize distributions such as tweedie or negative binomial distribution, as done by Olkonemi et al. (2023). These distributions belong to a family of exponential distributions that are capable of handling bimodal data, which, in the case of fixations, is due to the distinct distributions of the zero and non-zero values. In the workshop, we will present our results, where we compare various exponential distributions in their ability of modelling fixations.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Candace M. van der Stelt, Sarah E. Wallace, Michael Walsh Dickey, Kimberley Smith, Tessa Warren

Title: Identifying Pressure Points in Reading Comprehension in Acquired Alexia

Abstract:

Introduction: Alexia, an acquired reading impairment resulting from neurological insult, is traditionally assessed by reading aloud single words [1], but manifests in difficulty with both oral reading and reading comprehension [2]. Although impaired word reading in alexia has been studied and modelled [3,4], there is a dire need for characterization and explanation of these co-morbid reading comprehension impairments. The Reading Systems Framework [5] proposes that impairments at key “pressure points”—weaknesses in word identification subprocesses, such as phonological decoding and/or lexical-semantic representations – result in comprehension breakdown. This project aimed to determine whether these psycholinguistic pressure points apply to comprehension outcomes in acquired alexia.

Methods: We conducted an exploratory analysis using retrospective data from twenty-two individuals with acquired alexia following left-hemisphere stroke. Participants completed three subtests of the Psycholinguistic Assessments of Language Processing in Aphasia [6] (Reading by Regularity, Reading by Frequency and Imageability, and Nonword Reading) and all ten subtests of the Reading Comprehension Battery for Aphasia-Second Edition (RCBA-2) [7], measuring word-, sentence-, and paragraph-level comprehension. Four logistic mixed effects models were used to test if accuracy of reading aloud nonwords (phonology), regular vs. exception words (phonology plus lexical-semantics vs. lexical-semantics), high- vs. low-imageability words (high vs. low lexical-semantic difficulty), and high- vs. low-frequency words (high vs. low lexical-semantic difficulty) could predict reading comprehension on the RCBA-2. Helmert coding was used to contrast levels of reading comprehension on the RCBA-2: 1) text-level (sentences and paragraphs) versus word-level and 2) sentence- versus paragraph-level.

Results: Text-level comprehension was reliably lower than word-level comprehension in all four models. There were significant positive main effects of nonword ($\beta=0.44$, $p<.001$), exception word ($\beta=0.52$, $p<.02$), and high-imageability word ($\beta=0.65$, $p<.03$) reading accuracy on overall reading comprehension accuracy. There were no significant interactions between any oral word reading accuracy measure (i.e., different psycholinguistic subprocesses) and contrasts of reading comprehension level.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Discussion: Our results indicate that both phonological and lexical-semantic subprocesses impact word- and text-level reading comprehension in individuals with acquired alexia, warranting further investigation. We plan to use information from this analysis to develop controlled experiments to examine how psycholinguistic features affect eye-movement measures for individuals with alexia during reading comprehension tasks.

References:

1. Friedman, R.B. (2002). Alexia. In **Encyclopedia of the Human Brain** (pp. 111-117). Academic Press.
2. Webster, J., Morris, J., & Howard, D. (2023). Reading comprehension in aphasia: the relationship between linguistic performance, personal perspective, and preferences. **Aphasiology**, 37(5), 785-801. <https://doi.org/10.1080/02687038.2022.2039999>
3. Plaut, D., McClelland, J.L., Seidenberg, M.S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. **Psychological Review**, 103(1), 56-115. <https://doi.org/10.1037/0033-295X.103.1.56>
4. Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. **Psychological Review**, 108, 204–256. doi:10.1037/0033295X.108.1.204
5. Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. **Scientific Studies of Reading**, 18(1), 22–37. <https://doi.org/10.1080/10888438.2013.827687>
6. Kay, J., Lesser, R., & Coltheart, M. (1996). Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction. **Aphasiology**, 10(2), 159-180. <https://doi.org/10.1080/02687039608248403>
7. LaPointe, L. & Horner, J. (1998). Reading Comprehension Battery of Aphasia – Second Edition (RCBA-2). Pro-Ed, Inc.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Posters Day 1

Author(s): Yi Wang, Kesheng Wang

Title: Managing Mind Wandering by Integrating EEG and Eyetracking

Abstract:

Mind wandering during reading happens when attention shifts away from the text to unrelated thoughts. This shift will negatively impact on reading comprehension and retention. One of the main discussion in recent researches in detection of Mind wandering is whether a mono-sensor (Eye tracking or EEG) research design is sufficient. This research shows that by combining EEG (electroencephalography) with eye tracking offers several advantages over using eye tracking alone, especially in the context of psycholinguistic research and detecting mind wandering. The integration of these two methods allows for a more comprehensive understanding of cognitive processes and provides a richer dataset for analysis. The combination also enables researchers to correlate the exact moment of neural activity with specific reading behaviours. This paper also demonstrates a framework where machine learning algorithms provide a deeper insights into the neural correlates of mind wandering identifying associative brainwave patterns with irregular eye movements.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Ian Simpson, David Saldana, Mila Vulchanova, Martina Micai

Title: Exploring Reading and Feedback-Driven Learning in Autism Spectrum Disorder: An Eye Movements Study

Abstract:

Background. The Wisconsin Card Sorting Task (WCST) is a widely used neuropsychological assessment tool used to assess cognitive flexibility. Individuals with Autism Spectrum Disorder (ASD) present challenges in cognitive flexibility and rule-shifting domains.

Methods. Twenty-one Spanish-speaking autistic children and adolescents, along with matched controls classified 128 target cards based one of three possible rules (color, number, shape). Verbal responses to three strategy-related questions were gathered. Participants were informed if their responses were correct/incorrect but were not told that the classification rule changes after 10 sequential correct responses (known as a “set”). Reading fluency and comprehension (TALE-2000), along with grammatical structural knowledge (CEG) were also assessed.

Results. Behavioral and eye movement data revealed several significant differences between the two groups. Autistic individuals completed fewer sets, made a significantly higher percentage of total errors and non-perseverative errors, and provided fewer conceptual level responses compared to the control group. Both groups exhibited a significant increase in number of fixations and dwell time after receiving feedback, but the increase in both measures for incorrect responses, compared to correct responses was significantly greater for the controls. Reading comprehension correlated with response duration for perseverative errors solely within the control group. Conversely, within the ASD group, numerous correlations have been observed between reading fixation count, saccade amplitude, and dwell time concerning reading velocity.

Discussion. These findings suggest that individuals with autism encounter difficulties in error monitoring and response inhibition, potentially resulting in challenges in adapting to changing task demands and maintaining goal-directed behavior. Moreover, it appears that autistic individuals may not engage as extensively in error analysis and correction processes compared to the control group. For the controls, the time taken to respond during perseverative errors was associated with their overall ability to comprehend written material. Conversely, in autistic individuals, reading speed rather than overall comprehension ability was associated with their eye movement behavior during WCST. This underscores the necessity for targeted interventions aimed at enhancing feedback processing and promoting adaptive learning strategies, while also highlighting the different cognitive processes involved in reading tasks between autistic individuals and typically developing controls.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Adrielli Tina Lopes Rego, Martijn Meeter

Title: Simulating Regressions in Reading with Surprisal and Relative Word Importance from Language Models

Abstract:

What triggers regressions and what determines its destination during reading? This question remains challenging, due to the great variety of backward saccades and the number of competing explanations. One prevailing hypothesis is that regressions reflect comprehension difficulty, that is, readers go back in text to repair or reanalyze the textual input. Models of eye-movement control in reading have attempted to simulate regressions by triggering backwards saccades to the previous word when identification of the currently fixated word fails. This approach is limited to lexical processes and cannot account for long-range regressions. Here we investigate whether surprisal and relative word importance derived from language models (GPT2 and LLaMA) can respectively predict when backward saccades occur and to where. Surprisal is interpreted to reflect the difficulty of integrating the fixated word into the mental representation built from the previously read context. Relative word importance is a proxy of how relevant each word is in the previously read context to the prediction of the upcoming word. If these measures are indicative of regressions, this would be evidence of the link between regression and repair, especially associated with error costs from predictive processing. Additionally, the influence of saccade size is analyzed to investigate whether the surprisal and relative word importance simply reflect short-range, and thus local processing, or they can account for higher-level, long-range integration. Successfully predicting regressions in reading may advance our understanding about the interaction between oculomotor behavior and reading comprehension.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Eva Pavlinušić Vilus, Ana Matić Škorić, Kristina Cergol, Marijan Palmović

Title: The Role of Working Memory Span and Cognitive Flexibility in the Processing of Locality and Surprisal in Croatian: An Eye-Tracking Study

Abstract:

The ease with which we read and process words in a sentence depends on the number of factors intrinsic to the individual words it is comprised of, such as length and frequency, but also on some sentence-level factors, e.g., locality and the predictive value of the preceding context. Certain intraindividual traits are also known to impact the efficiency of language processing while reading, namely the working memory (WM) span and cognitive flexibility (CF) (Cartwright, 2007).

The aim of the present study is to examine the role of WM and CF in the processing of sentences containing words that vary in contextual predictability and distance from their syntactic heads, i.e., in surprisal and locality. This is an on-going eye-tracking study, currently in the stimuli pre-testing stage, that is to the best of the authors' knowledge original in the Croatian context.

Thirty native speakers of Croatian have performed a sentence completion task designed to validate the low-surprisal sentence items. Based on the results of this pretest, the final list of stimuli with 20 sentences per condition will be prepared, together with an equal number of filler sentences. Two factors, locality and surprisal, will be manipulated in a 2x2 experimental design in a single-sentence reading experiment. Native speakers of Croatian will participate in the core experiment. Eye-movements will be recorded at a sampling rate of 1000 Hz using the SR Eyelink Portable Duo eye-tracker. Participants will complete a WM test battery (Lewandowsky et al., 2010) and a computerized version of Berg's card sorting task (Grant & Berg, 1948) for the assessment of their CF.

Both early and late measures will be analyzed, e.g., first fixation durations, average fixation durations, and total reading times. Results are expected to reflect the patterns observed in other languages, while the effect of distance and surprisal will be discussed with respect to relevant typological characteristics of Croatian. Moreover, participants with shorter WM span are expected to exhibit more difficulties with the processing of long-distance dependencies, whereas higher CF is expected to enable faster processing of high-surprisal words.

REFERENCES

Cartwright, K. B. (2007). The Role of Cognitive Flexibility in Reading Comprehension: Past, Present, and Future. In Israel, S. E. & Duffy, G. (Eds.), *Handbook of research on reading comprehension* (pp. 115–139). New York: Routledge.

Grant, D.A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in Weigl-type card-sorting problem. *Journal of Experimental Psychology* 38, 404-411.

Lewandowsky, S., Oberauer, K., Yang, LX. et al. (2010). A working memory test battery for MATLAB. *Behavior Research Methods* 42, 571–585. <https://doi.org/10.3758/BRM.42.2.571>



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Jacquelyn F. Stochel, Candace M. van der Stelt, Alexander Swiderski, Michael Walsh, Sarah E. Wallace, Tessa Warren, Kelly Knollman-Porter, Karen Hux

Title: Examining the Effect of Word Surprisal on Eye Movements During Story Reading in Neurotypical Adults and People with Aphasia

Abstract:

Over two-thirds of people with aphasia have reading impairments impacting their quality of life. This justifies detailed investigation of reading to improve assessments and treatments that support this population. The present study used existing eye-tracking data from nine neurotypical adults and nine people with aphasia to evaluate the relationship between surprisal (i.e., word unexpectedness in context) and eye-movement measures during story reading.

Surprisal was estimated with GPT-2 by computing the log probability of each story noun given its preceding context. Four eye-tracking measures were assessed for each noun: probability of first pass fixation (PF), probability of first pass regression (PR), gaze duration (GD), and total fixation duration (TFD). Linear mixed effects models were employed to investigate the relationship between these eye tracking measures and group (aphasia, neurotypical), noun surprisal, and their interaction, while controlling for lexical variables (e.g., frequency) and random variables (e.g., person).

People with aphasia had significantly higher GD and TFD compared to neurotypical adults. Surprisal also reliably predicted all four eye-tracking measures, where higher surprisal nouns were associated with higher PF but lower PR, GD, and TFD. There was a marginal interaction of group and surprisal in PF, such that people with aphasia were even more likely to fixate on high surprisal nouns than low surprisal nouns compared to neurotypicals. The remaining models revealed no other significant interactions between group and surprisal. Exploratory analysis also revealed that the semantic distance between the story noun and a predicted alternative by GPT-2 affected GD and TFD.

The finding that higher surprisal nouns were more likely to be fixated was expected, but the direction of surprisal effects on other measures was unexpected. Future analyses will examine whether word position effects may explain these unexpected findings. The surprisal effects and lack of interaction effects were also consistent with previous findings of spared predictive processing in people with aphasia. Our lab plans to conduct future research on the relationship between surprisal, eye movements, and reading comprehension to elucidate the role of surprisal in reading assessments, and how surprisal modulation can be leveraged in the treatment of reading impairments in aphasia.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Jiaxin Li, Vera Demberg

Title: Modelling Individual Differences in Locality Effects with Lossy-Context Surprisal

Abstract:

Locality effects, slowdowns produced by long dependencies, are generally assumed to reflect limited working memory. Puzzlingly, locality effects are sometimes stronger in readers with higher working memory capacity (WMC) [1] — at odds with the idea that they reflect limited memory. Here, we interpreted these puzzling findings in terms of lossy-context surprisal (LCS, [2]), which assumed that human processing difficulty reflected surprisal computed from incomplete context representations. We hypothesised that people with different amounts of WMC were modelled by LCS with different amounts of remembered context.

We considered data from two experiments targeting locality effects, in Spanish [1, SPR] and English [3, Maze]. We computed lossy-context surprisal by truncating the context [4] for Spanish, and the resource-rational model of [3] for English.

In both datasets, we found the same pattern: when the available context was extremely limited or very large, LCS predicted no locality effects. It was only in between, with intermediate amounts of context, that locality effects were predicted.

In the Spanish dataset, human RTs on the critical region were matched by surprisal computed from 1 (low WMC) or 2 (high WMC) context words, similar to the range found to work best by [4]. Locality effects were predicted at 2 words, and not at 1 words of context, explaining the counterintuitive finding of [1].

In the English dataset, as a proxy for WMC, we first measured how well each subject's RTs on the fillers were modelled by surprisal with low or high forgetting rates. On the critical region, subjects whose filler RTs were better predicted by higher forgetting rates also showed a stronger locality effect on the critical condition. The direction of the effect was thus opposite to the Spanish data, where low-WMC subjects showed smaller locality effects. Indeed, locality effects predicted by LCS showed this pattern at the forgetting rates most predictive of filler RTs, reconciling the contradicting findings in the two datasets.

[1] Nicenboim et al., 2016, *Frontiers in Psychology*

[2] Futrell et al 2020, *Cognitive Science*

[3] Hahn et al. 2022, *PNAS*

[4] Kuribayashi et al. 2022, *EMNLP*



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Pei-Chun Chao, Jou-An Chung, Jie-Li Tsai, Chia-Ying Lee

Title: Assessing Reading Comprehension in Chinese Dyslexia and ADHD: Insights from Eye Movements and Fixation-Related Potentials in Natural Reading.

Abstract:

Dyslexia and Attention Deficit/Hyperactivity Disorder (ADD/ADHD) are two of the most prevalent neurodevelopmental disorders. Despite both encountering obstacles in reading comprehension, the underlying issues stem from distinct cognitive deficiencies. Dyslexic individuals struggle with phonological processing and word decoding, while those with ADHD exhibit impairments in executive function. Current assessments rely heavily on subjective observation and standardized tests, often leading to confusion in diagnosis. To address this, our study aimed to objectively evaluate online reading comprehension in adults with dyslexia and ADD/ADHD using simultaneous recording of eye movements and fixation-related potentials (EMFRPs). We established a normative EMFRP database with forty-seven adults while reading traditional Chinese articles. We applied linear mixed-effects models (LMMs) for single-trial FRPs to examine word frequency-predictability interaction effects (WF*WP) on the mean N400 amplitudes between 375-475 ms in central-posterior regions. The mean amplitude decreased with increasing predictability only for infrequent words, indicating that most people can effectively utilize contextual information to understand unfamiliar words. Subsequently, the EMFRP data were collected in five adults with dyslexia (Case 1), ADD (Case 2 & Case 3), and ADHD (Case 4 & 5). Compared to the norm, Case 1 read articles with shorter total viewing times (TVTs) but more regressions and responded with less accuracy and average response times (RTs). The WF*WP effects on N400s occurred in frontal regions between 300-400 ms. Case 2 read articles with shorter TVT and responded with average accuracy and shorter RTs. The WF*WP effects on N400s occurred in posterior regions between 425-525 ms. Case 3 read articles with longer TVTs and more regression and responded with average accuracy but longer RT. Only WF effects on N400s were found. Case 4 read articles with longer TVTs and more regression and responded with less accuracy and average RTs. Only WF effects on N400s were found. Case 5 read articles with shorter TVT and less regression and responded with average accuracy and RTs. Only the main effects of WF and WP on N400s were found. The comprehensive EMFRP profiles could reveal unique reading mechanisms and assess reading proficiency for each individual with dyslexia or ADD/ADHD.



MultipleYE

Author(s): The MultipleYE team

Title: The MultipleYE corpus

Abstract:

Learn more about the MultipleYE data collection — a multi-lab effort to collect a multi-lingual eyetracking-while-reading corpus.

The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Posters Day 2

Author(s): Jou-An Chung, Pei-Chun Chao, Jie-Li Tsai, Chia-Ying Lee

Title: Effect of Working Memory Load in Natural Sentence Reading in Adolescents with ADHD and ADD

Abstract:

Individuals with ADHD or ADD exhibit poor reading comprehension which renders the academic underachievement throughout school age. Prior studies reported that executive dysfunctions such as deficits in working memory (WM) may explicate their reading difficulty. Nevertheless, the extant evidence remained equivocal since most findings relied on the correlation between WM capacity and performance on the offline standardized reading assessments. Hence, this study examined the effect of WM load on natural reading among ADHD, ADD and normal controls. Simultaneous recording of Eye-Movement (EM) and Fixation-Related ERPs (FRPs) were applied to provide complementary measures to capture cognitive operations in natural reading.

Direct manipulation of WM-load was determined by the varying filler-gap distance. Low WM-load condition comprises sentences with the simplest Chinese Subjective Relative Clause ([GAP] RCverb (協助/help) + RCnoun (居民/resident) + 的(DE) + Head noun(HN) [Filler] (志工/volunteer)). As for medium and high-WM load conditions, one constituent- adverb (ex: 熱心/enthusiastically) and two constituents- a locative preposition and an adverb (ex: (在社區/in community) + (熱心/enthusiastically)) were added prior to RCverb, respectively. Fifty-three native Chinese speakers without ADHD, ADD or Dyslexia read 81 sentences while EMFRPs were recorded. EM patterns on HN where the ambiguity is resolved were compared among conditions. LMM analysis revealed the significant main effect of WM load on first-pass (first fixation duration) and second-pass measures (rereading rate and regression-in rate), indicating the effect of WM load on word recognition and sentence integration.

Thus far, six ADHD and six ADD completed the same task. Preliminary results on overall EM patterns revealed that ADHD showed the longest total reading time and the greatest number of fixations. Compared to normal controls, ADHD and ADD initiated more backward saccades as WM load increased. However, ADHD generated long and unstable backward saccades; whereas, ADD made short backward saccades. Moreover, ADHD exhibited longer forward saccades and riskier reading scan path. Current findings demonstrated that modulations of WM load on sentence comprehension can differentiate three groups. EM patterns and FRPs can optimize the course of differential diagnosis.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Oksana Ivchenko, Natalia Grabar

Title: Predictive Modelling of Fixations for Deciphering Reader Engagement

Abstract:

Reading comprehension is a complex cognitive process crucial for both academic and everyday information processing (1, 2). This research project aims to study the reading of different text types through eye movement measures, particularly focusing on predicting the number and duration of fixations within texts. Our dataset includes four types of texts: medicine-related articles from Wikipedia, medical cases (3), their simplified versions, and texts on general topics (4). We exploit eye-tracking data collected from 10 participants using the Tobii Pro Spectrum at 600 Hz. These participants are all native French speakers, aged between 18 and 55 years, without any medical education. Their eye-tracking data shows high-accuracy calibration, thus facilitating the association of textual features with the corresponding eye-tracking metrics.

We performed statistical analysis of the data, which revealed several observations: 1. Eye-tracking characteristics do not exhibit a normal distribution; 2. There is a significant difference in reading patterns across these four types of texts; 3. A high correlation was observed between the duration of fixations and the number of fixations on a word, indicating that words with longer fixations tend also to have more fixations.

Then, starting from these eye-tracking data, our purpose is to generate language models for the prediction of eye-tracking metrics, like the number and duration of fixations. Our hypothesis is that the correlation between textual characteristics and eye movement patterns, can also be predicted by language models. Hence, our objective is to automatically decipher the complex relationships between text and reader engagement. Our method consists of several steps: verification of alignment between words and the corresponding eye-tracking metrics, tokenization of text data, conversion of raw sentences into a structured format that the model can interpret, enrichment of the representation with language information (frequency, length, etc. of words). Building upon a pre-trained FlauBERT base (5), we have developed a custom neural network model equipped with a regression layer designed to predict eye-tracking features directly from text inputs. We will present the results obtained from this model and discuss them.

Bibliographical References

Judith Wylie, Jennifer Thomson, Paavo Leppanen, Rakefet Ackerman, Laura Kanninen, and Tanja Prieler. 2018. Chapter 3. Cognitive processes and digital reading, pages 57–90.

Ooi K. and Khaidzir Ismail. 2011. The relationship between cognitive processing and reading. *Asian Social Science*, 7.

N. Grabar, Clément Dalloux, and V. Claveau. 2020. Cas: corpus of clinical cases in french. *Journal of Biomedical Semantics*, 11.

Natalia Grabar and Rémi Cardon. 2018. CLEAR – simple corpus for medical French. In *Proceedings of the 1st Workshop on Automatic Text Adaptation (ATA)*, pages 3–9, Tilburg, the Netherlands. Association for Computational Linguistics.

Hang Le, Loïc Vial, Jibril Frej, Vincent Segonne, Maximin Coavoux, Benjamin Lecouteux, Alexandre Allauzen, Benoit Crabbé, Laurent Besacier, and Didier Schwab. 2020. FlauBERT: Unsupervised Language Model Pre-training for French. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 2479–2490, Marseille, France. European Language Resources Association.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Francesca Zermiani, Prajit Dhar, Ekta Sood, Fabian Kögel, Andreas Bulling, Maria Wirzberger

Title: InteRead: An Eye Tracking Dataset of Interrupted Reading

Abstract:

Eye movements during reading offer a window into cognitive processes and language comprehension, but the scarcity of reading data with interruptions -- which learners frequently encounter in their everyday learning environments -- hampers advances in the development of intelligent learning technologies. To this end, we recently introduced InteRead (Zermiani et al. (2024)) -- a novel 50-participant dataset of gaze data recorded during self-paced reading of real-world text. InteRead further offers fine-grained annotations of interruptions interspersed throughout the text as well as resumption lags incurred by these interruptions. Our analyses show that our dataset is in line with relevant previous findings (Schilling et al. (1998), Kiegl et al. (2004)), particularly regarding first pass fixations. Besides, we find that the logarithmic token frequency exhibits a non-linear relationship with word length, as found in Kiegl et al. (2004).

In this Workshop, we plan on presenting our findings as well as starting a discussion on how to handle zero-count fixation data. Fixations, measured either as durations or raw counts, follow an extremely skewed distribution. In our experiments, we found 59% of the fixation counts to hold a value of zero, with the distribution being bimodal. While the standard approach in dealing with zero-count fixations entails setting a lower-bound, for instance discarding fixation durations less than 100ms, we propose to retain the fixation data as it is, but instead utilize distributions such as tweedie or negative binomial distribution, as done by Olkonemi et al. (2023). These distributions belong to a family of exponential distributions that are capable of handling bimodal data, which, in the case of fixations, is due to the distinct distributions of the zero and non-zero values. In the workshop, we will present our results, where we compare various exponential distributions in their ability of modelling fixations.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Van Rynald T Licalalde, Tessa Warren, Ariel N James, Duane G Watson

Title: Online Eye-Tracking and Self-Paced Reading Measures are Only Portable at the Item Level

Abstract:

Eye-tracking and self-paced reading (SPR) are both used to study sentence processing and reading comprehension. Psycholinguists often predict that findings in one will hold in the other. This assumes that eye-tracking and SPR measures are *portable*, i.e., that these measures directly map to each other. We evaluate this assumption here. Specifically, we ask at *what sampling level* (observation, item, subject) and at *what unit of aggregation* (grand average, averaging within syntactic region, averaging within word, no averaging) are eye-tracking and SPR measures portable.

116+ subjects read the same 120 English sentences in the same order in separate eye-tracking and SPR sessions that were 7+ days apart. Eighty sentences included a subject/object-extracted relative clause (RC) manipulation (1); 40 sentences included a lexical age-of-acquisition (AoA) manipulation (2).

1. The guards **who provoked the spectators [who the spectators provoked]** before the protest increased security around the area.
2. Jessica spotted the **dark branch [orchid]** before the bright flowers.

First, we adjusted the SPR RTs and eye-tracking measures for overall processing speed. We then estimated portability by correlating the adjusted RTs with their matching adjusted eye-tracking measures (e.g., gaze duration, go-past, total time) at various units of aggregation for each target sampling level. For example, to assess portability for word averages at the item-level, we computed the mean adjusted RT and mean adjusted eye-tracking measures for each word within each item and correlated these means across items.

For the RC items, we only observed portability for region averages at the item-level ($r = [0.38, 0.78]$). For the AoA items, we only observed portability for word averages at the item-level ($r = [0.35, 0.76]$). These results support item-level portability: we can predict result *****R***** on sentence **S** in eye-tracking from observing **R** on **S** in SPR. But given that we did not find subject-level or observation-level portability, we are not justified in mapping a person's behavior or single observations between methods. These results indicate that researchers should be careful not to assume that people's performance will be equivalent across methods to the extent that these methods are used to measure online reading comprehension.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Mario Romero, Clara Soberats, Ladislao Salmerón

Title: Pay Attention to the Text: Do Eye-Movements Reflect Changes of an Intervention to Boost Long Text Reading in Adolescents?

Abstract:

The shallowing hypothesis, the idea that as readers interact with digital media, they develop a mindset favoring rapid, short, and immediate gratification consumption of information, may explain comprehension difficulties in long digital texts (Annisette & Lafreniere, 2017). Interventions focusing on improving attention to the text seem promising for enhancing adolescents' comprehension. A major challenge to assess the effectiveness of such interventions is to measure attention during reading long texts, without relying on subjective ratings reading. In this study, we explore the extent to which eye-movements could be used for such purpose.

Twenty-two adolescents (12 girls, 14 years) from a semi-private high school participated in a 16-week reading intervention on an online platform. One week before and after the intervention, participants read four expository texts on biology and history (two at each session), while their eye-movements were recorded. Texts were composed of 11 paragraphs, and had approximately 1.500 words.

Eye movement data was recorded using a LogicOne LD250 (Eyelogic GmbH, Berlin, Germany), a video-based remote system equipped with dark pupil technology and a sampling frequency of 250 Hz. We used paragraphs as areas of interest (AOI). We considered the number and sum of the fixation durations of each AOI, as well as average fixation duration. We used linear mixed-effects models (LMMs) with subject, text as random effects, and time, paragraph (1-11), and its interaction as fixed factors.

After the intervention students increased their overall time spend on texts, not by changing the number of fixations, but rather by increasing the average time of fixations. Regarding the effect of paragraph, taking as reference the initial paragraph, students fixated less and for less time early in the text, as well as late in the text. Finally, the interaction between time and paragraph indicated that, at pre-test, average fixation duration decreased early and late in the text. By contrast, at post-test, average fixation duration increased early and late in the text.

These findings support the use of eye-movements as objective and unobtrusive indicators of students' attention during reading, as they capture both effects within long texts, as well as effects of a long intervention.



The 2nd Workshop on Eye Movements and the Assessment of Reading Comprehension

Author(s): Nadina Gómez-Merino, Antonio Ferrer, Ana García-Blanco, Inmaculada Fajardo

Title: Idiom Comprehension in Autism Spectrum Disorder: Evidence from Eye Movement Measures of Reading

Abstract:

Previous research has found that individuals with autism spectrum disorder (ASD) show poorer comprehension of figurative language than peers with typically development highlighting that the type of figurative Language tested (e.g. irony, metaphors or idioms) was significantly related to differences in effect sizes (Kalandadze et al., 2018). In this study, we go deeper in investigating how participants with ASD process idioms, a particular type of figurative language (complex multiword linguistic expressions that convey a meaning not derived from the combination of the individual literal meanings of its consisting words (e.g. Titone & Connine, 1994) in both oral and written form. Nineteen ASD participants (age range: 12 to 16) without intellectual disability (Mean verbal IQ= 106; Mean non-verbal IQ=113) were asked to read/listen sentences with very familiar idioms (previously piloted) that could either be interpreted as figurative or non-figurative depending on the previous context of the sentence (e.g. break the ice). Afterwards, they were asked idiom comprehension questions. Accuracy, question reaction time and eye-movements during sentence reading were registered.

Linear and generalized mixed models analyses showed no effects of presentation format (oral vs. Written) and idiomaticity (literal vs. figurative) on accuracy (above 90% across conditions) but we found an interaction between format and idiomaticity for question reaction time: ASD participants took longer to answer comprehension questions in the literal than in the figurative condition but only in the written format. Regarding eye-movement (only registered for written format), ASD participants showed significantly longer first pass time, total time and regression path time in the idiom area when it was preceded by a figurative context than by a literal one. In summary, we conclude that ASD participants read idiomatic expressions more cautiously than literal ones, which paradoxically might have made them respond more quickly to comprehension questions in the figurative condition. Complementary analysis of online (during reading) and offline (after reading) measures suggests that eye movements and other types of online measures (e.g., EEG measures) are essential for drawing conclusions about figurative language processing in ASD individuals and theoretical postulates about idiom processing (direct retrieval vs. compositional accounts, Titone et al., 2019).