Study of Medical Text Reading and Comprehension Through Eye-Tracking Features

Oksana Ivchenko
PhD Student
(under supervision of Natalia Grabar)

UMR 8163 STL CNRS, Université de Lille

Louvain-La-Neuve, 15/11/2024

Objectives

- Making information accessible to:
  - Children or adults with reading difficulties, foreigners, adults with neurocognitive disorders, **general public (specialized texts)**
- Preserving their original meaning
- Application of suitable simplification and rewriting rules

Challenges

- Excessive standardization leading to loss of precision
  - *use common words, short sentences, do not use negation...*
- Difficulty in interpreting and adapting rules
- Inefficiency in complex contexts
- *Propose specific methods for the detection and characterization of* **reading difficulties**

- **Reading** is a crucial component of cognitive processes
- A key conduit for individuals to acquire information (Aarsland et al. [2021], K. and Ismail [2011])

Comprehension of information can pose challenges (Khoong et al. [2019], Novakova et al. [2023]):

- Education level
- Language proficiency
- Health conditions
- ...

In neurodegenerative disorders:

- Reading and comprehension are significantly impeded (Ekstrand et al. [2021])

Objective: Study the reading easiness with eye-tracking

Importance of the Study on Medical Texts

- **Specialized Vocabulary:**
  - Use of technical terms unfamiliar to non-experts
  - Example: *akathisia* (inability to remain seated)
- **Dense Information:**
  - Heavy information load packed into each sentence or paragraph
- **Long Sentences:**
  - Use of longer and more complex sentence structures
- **Abstract Ideas:**
  - Can be challenging to understand without specific training

Eye-tracking:
Objective insights into human behavior across various fields

Eye-tracking offers a window into visual attention and cognitive processes:

- Measure focus points and eye movement paths
- Quantify fixations to understand engagement

Widely applied in:

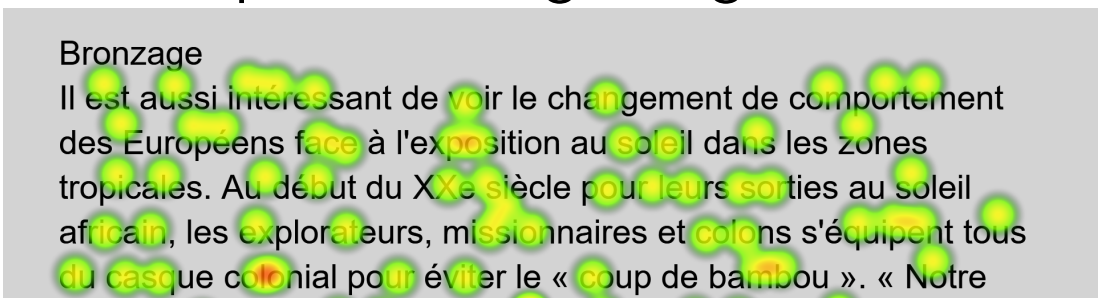
- **Psychology and Medical Research:**
  - Parkinson's disease (Archibald et al. [2013])
  - Alzheimer's disease (Boucart et al. [2014])
  - Stroke (Walle et al. [2019])
- **Market Research and User Experience (UX) Design** (García and Cano [2022])

Use of eye-tracking in **Psycholinguistics** and **Linguistics:**

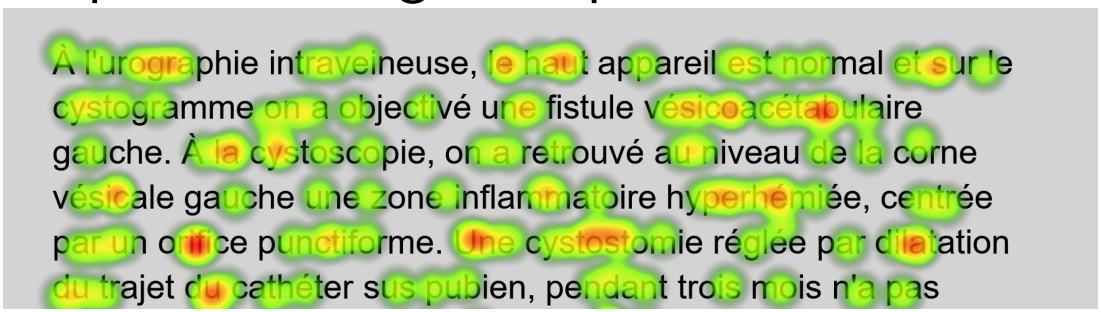
- Second language acquisition (Soroli et al. [2012], Godfroid et al. [2020])
- Reading Comprehension and Processing (Kuperman et al. [2018], Kim et al. [2019], Mézière et al. [2023])
- Semantics and Pragmatics (Magnabosco and Hauk [2023], Salicchi et al. [2023])
- Phonology and Prosody (Marquis et al. [2020], Ronai et al. [2019])
- ...

Fixations are brief pauses for information processing during reading
Longer duration of fixations indicates some difficulty in processing of the information or a specific interest of the readers

Example of reading of a general text



Example of reading of a specialized medical text



Video Example

Three steps of the research work:

1. Collection of eye-tracking data
2. Statistical analysis of eye-tracking data
3. Automatic prediction of eye-tracking indicators

Objectives:

- Collect eye tracking data on different types of texts
- Analyze reader engagement with medical / non-medical texts
- Understand processing differences among various reader groups

Corpus

- **Sources of texts in French:**
  - Wikipedia articles: general and medical topics
  - clinical cases
- **Texts Compiled** (12k occ.):
  - 6 medical texts (Stroke, Autopsy, Chikungunya, Ulcer, etc.)
  - 4 general texts (Weekend, Quince, Popcorn, Camelot)
  - 4 clinical cases (detailed accounts on patients including symptoms, diagnosis, treatment, and follow-up)
  - simplified versions of these texts (13k occ.) - manual syntactic and lexical simplification

Design of Comprehension Statements

- To evaluate participant understanding while maintaining natural reading conditions
- Statements are strategically placed only on some text segments to keep the reading as natural as possible
- One or two statements related to the segment just read
- Answer Options: *True*, *False*, or *I don't know*

Each participant read one of the text versions: original or simplified

Example of a Comprehension Statement

O: *His hemodynamics were stable and there were no signs of shock*

S: *The blood circulation (hemodynamics) was stable. The blood pressure was normal*

- **Statement:** *The patient's blood circulation was regular.*
- Correct Answer: *True*

Pre-Experiment Setup

- Questionnaire: Daily understanding of medical information
- The first text is the reference text, with two comprehension questions
- The corpus was carefully divided into two distinct Sets to ensure a balanced representation of text types in each set
- **Version A:**
  - Text 1: Original form
  - Text 2: Simplified form
  - Text 3: Original form
  - etc.
- **Version B:**
  - Text 1: Simplified form
  - Text 2: Original form
  - Text 3: Simplified form
  - etc.

Equipment: Tobii Pro Spectrum eye-tracker, operating at 600 Hz

Current Participant Data

- Total participants: 82 (+3 for pre-test)
- Age range: 18 to 55 years
- Backgrounds: students, PhD candidates, Postdoc, employed individuals, speech therapist students

(Medical) Language Proficiency

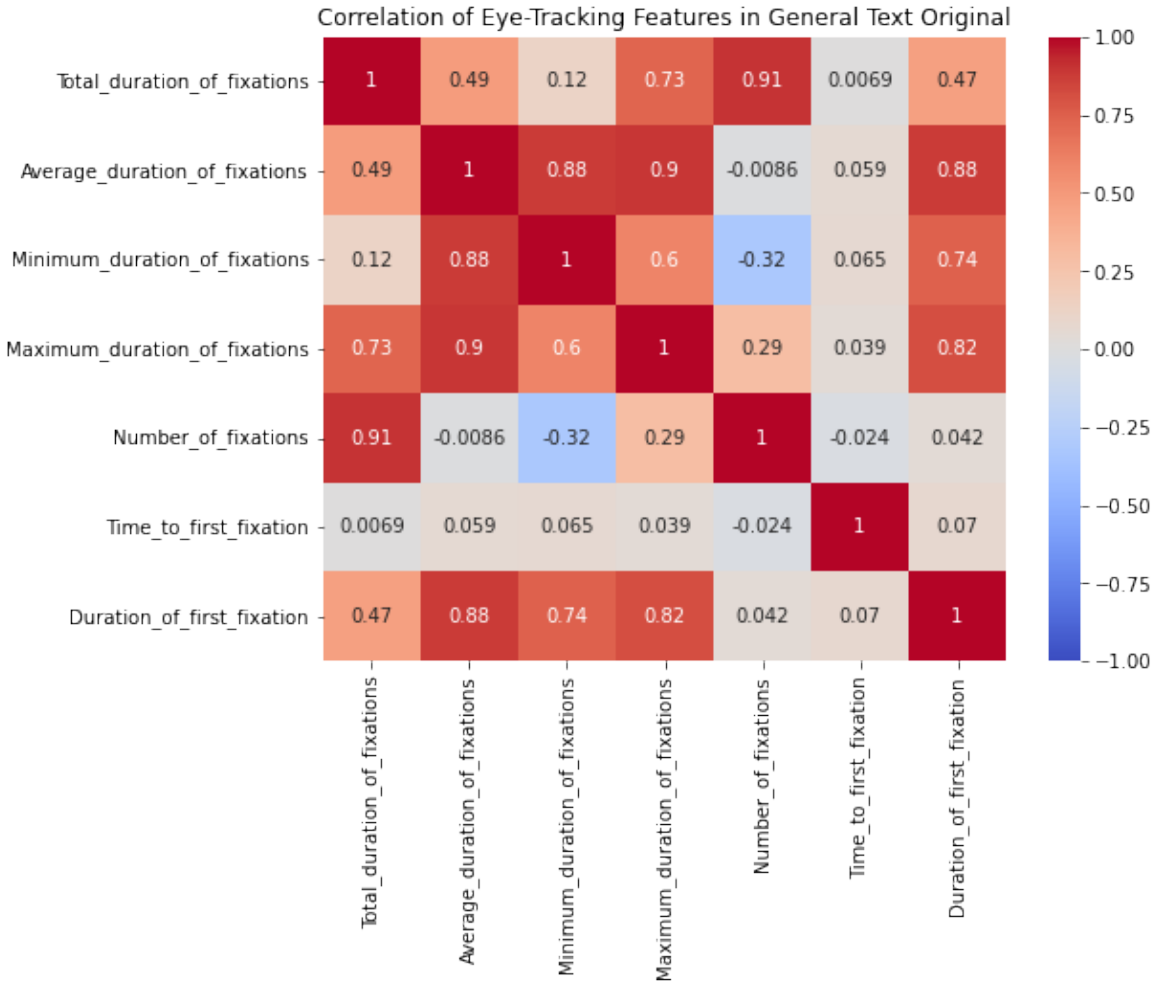
- *Native French speakers: 52*
- Non-native French speakers (C1+): 11
- *Speech therapists: 19*

Focus: Expertise in communication disorders, swallowing, orofacial motor skills, and cognitive functions, including care for patients with cerebral injuries

Three steps of the research work:

1. Collection of eye tracking data
2. Statistical analysis of eye-tracking data
3. Automatic prediction of eye-tracking indicators

Correlation between Eye-Tracking indicators

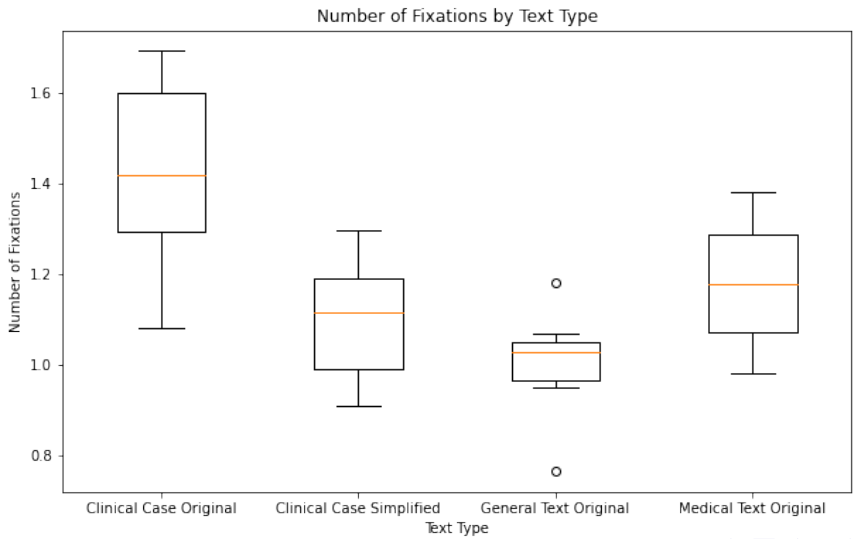


Study Focus

- Analysis based on eye movement data from 12 participants

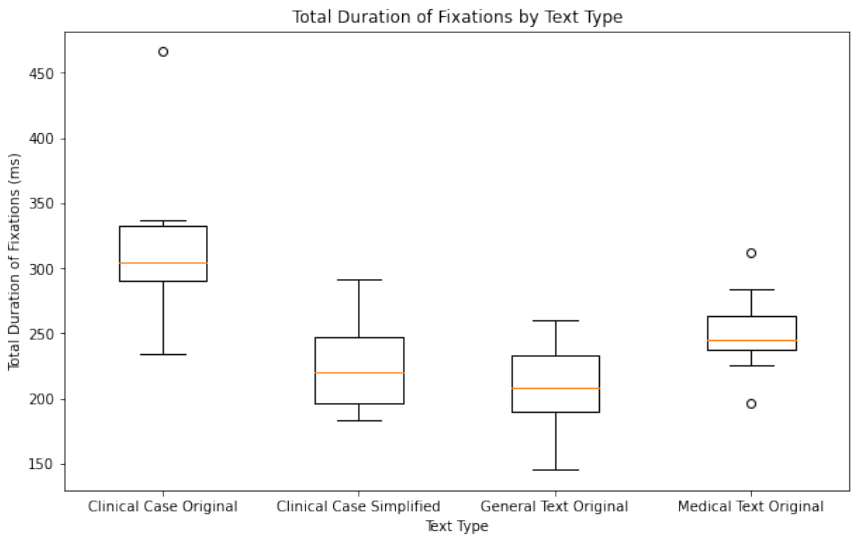
Average Number of Fixations per Word

Text Type	Count	Mean	Std	Max
Clinical Case Original	3184	1.42	1.73	23
Clinical Case Simplified	3024	1.11	1.18	13
General Text Original	7224	1.00	0.98	9
Medical Text Original	23895	1.17	1.28	24



Average Total Duration of Fixations Across Text Types (ms)

Text Type	Count	Mean	Std	Max
Clinical Case Original	3184	319.58	437.35	6397
Clinical Case Simplified	3024	227.50	261.22	3600
General Text Original	7224	207.59	226.08	2563
Medical Text Original	23895	250.54	300.27	5053



Three steps of the research work:

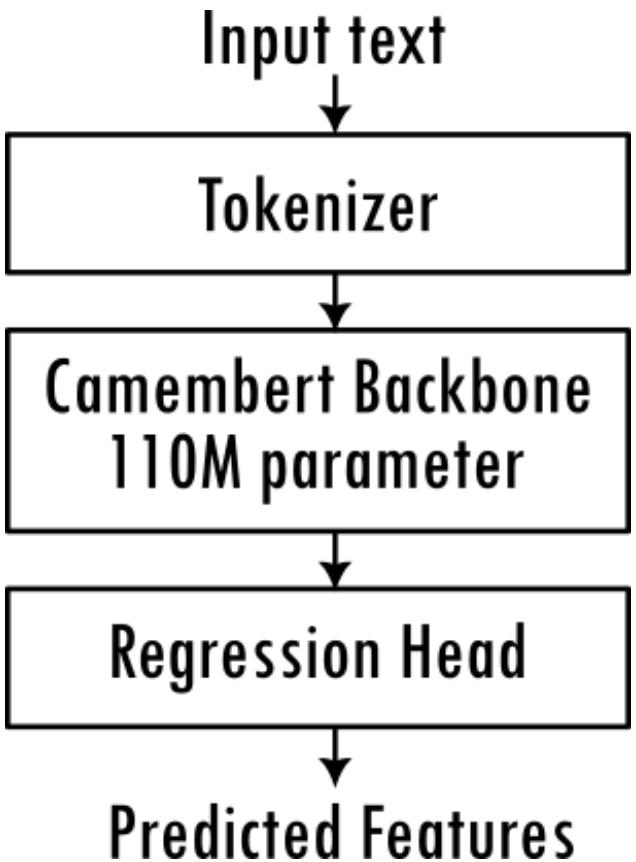
1. Collection of eye tracking data
2. Statistical analysis of eye-tracking data
3. Automatic prediction of eye-tracking indicators

Objective

- Purpose: Train AI language models to predict eye-tracking metrics across text types
  - number and duration of fixations
- Hypothesis: Eye movement patterns can be predicted from text data using deep-learning LMs

Model Development

- Base: Pre-trained CamemBERT model (Martin et al., 2020), a transformer LM trained on French texts
- Customization: Add a regression layer and fine-tune the model to predict eye-tracking features from text inputs



Input to CamemBERT:

- word tokenization
- assignment of the corresponding eye-tracking features to all tokens of words
- on example of 'Érythème':

token	*duration*	*fixations*
_É	0.648	1
_ry	0.648	1
_thème	0.648	1

Texts: 29 texts from different genres, manually simplified
Participants: 30 French-speaking general students
Age Range: 19 to 37 years

Text Type	**Nb Texts**	**Tokens**	**Sentences**
Medical Texts	5	4,976	722
General Texts	14	13,769	1,389
Clinical Cases	2	653	116
Simplified Medical Texts	6	6,941	1,396
Simplified Clinical Cases	2	805	193
Total for training*	**87**	**77,033**	**3,816**

**Each texts read by 1 to 4 participants*

Evaluation metrics:

- **Mean Squared Error (MSE)**: Average squared error; lower is better
- **Root Mean Squared Error (RMSE)**: Typical error size in original units
- **Mean Absolute Error (MAE)**: Average absolute error

$\Rightarrow$ **R^2 Score**: Proportion of variance explained; closer to 1 is a better model fit

- *No existing work for French*
- German (Salicchi et al. [2022]):
  - R^2 Score for predicting total reading time (TRT): 0.628
  - regression models using linguistic features

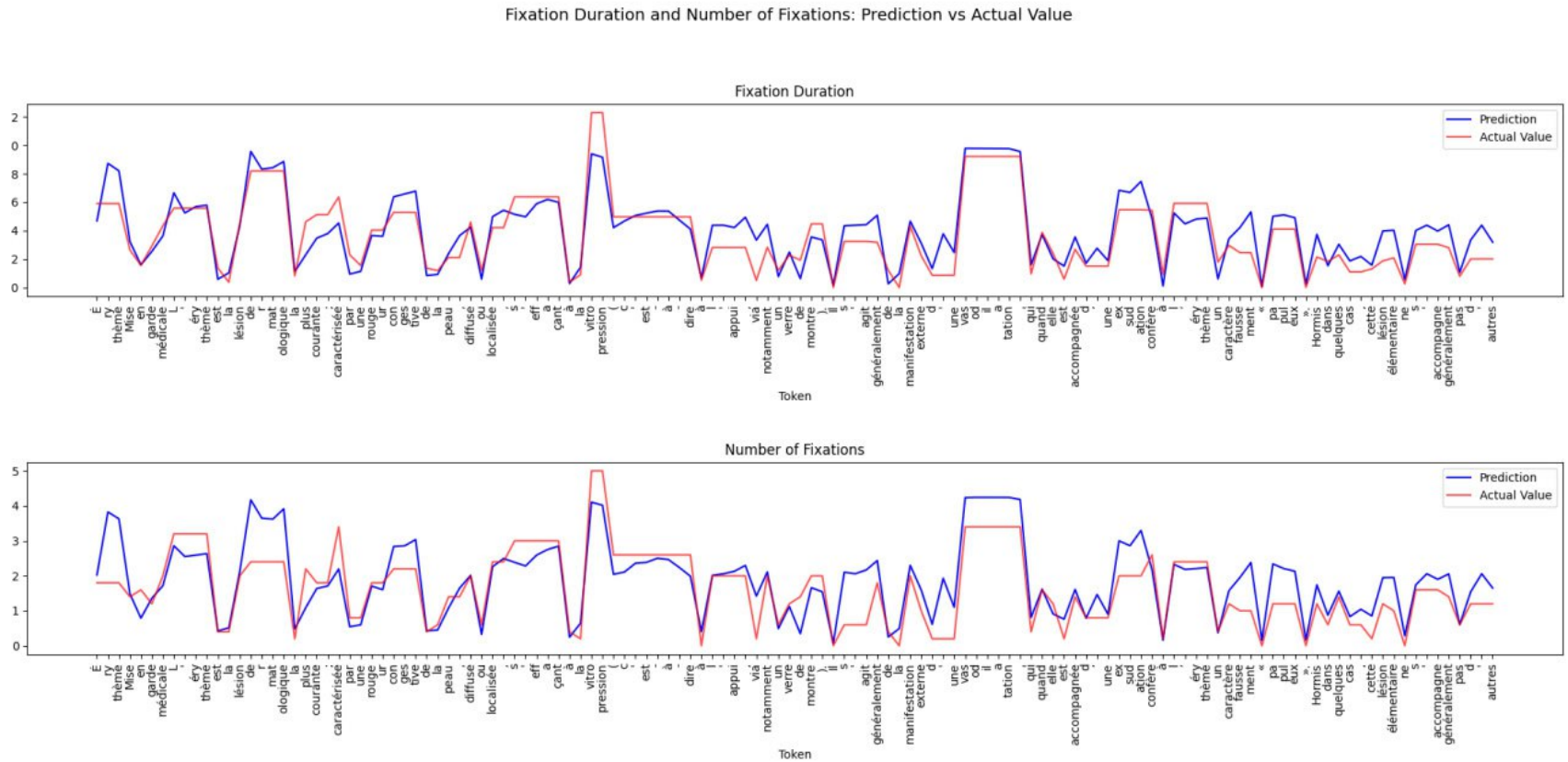
Average performance during the prediction step

Metrics	MSE	RMSE	MAE	R^2 Score
Number of Fixations	0.5572	0.7465	0.5395	0.6537
Fixation Duration (sec)	0.0395	0.1986	0.1266	0.6191

Features	Variance (σ^2)	Std Dev. (σ)
Number of Fixations	0.1036	0.3219
Fixation Duration (sec)	1.6089	1.2684

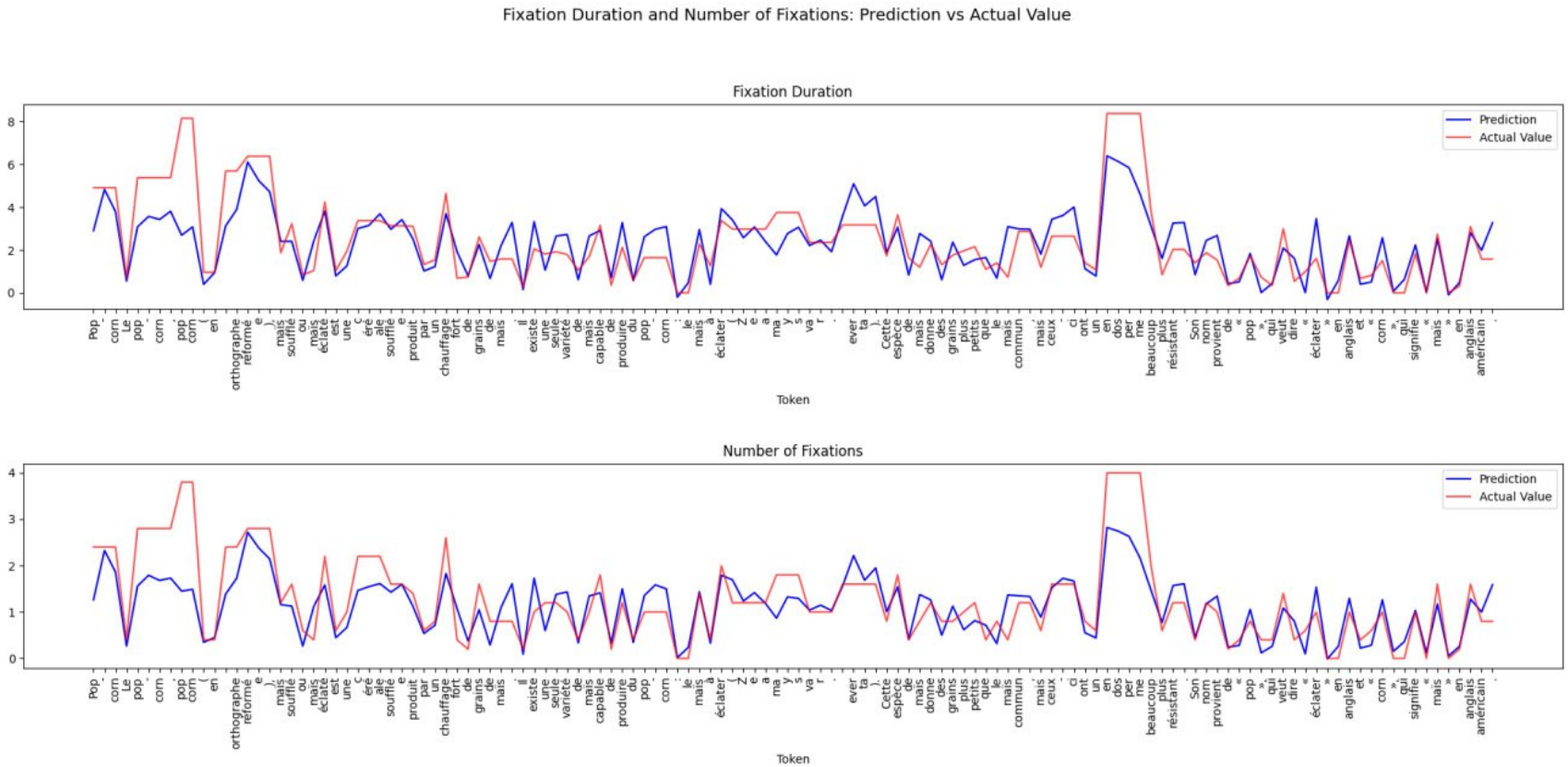
- Variability in Fixation Duration is higher
- Fixation Duration is predicted with higher accuracy (lower errors) compared to Number of Fixations
- The R^2 scores show the good performance of the model, but explains the variance slightly better for Number of Fixations

Medical Text Original



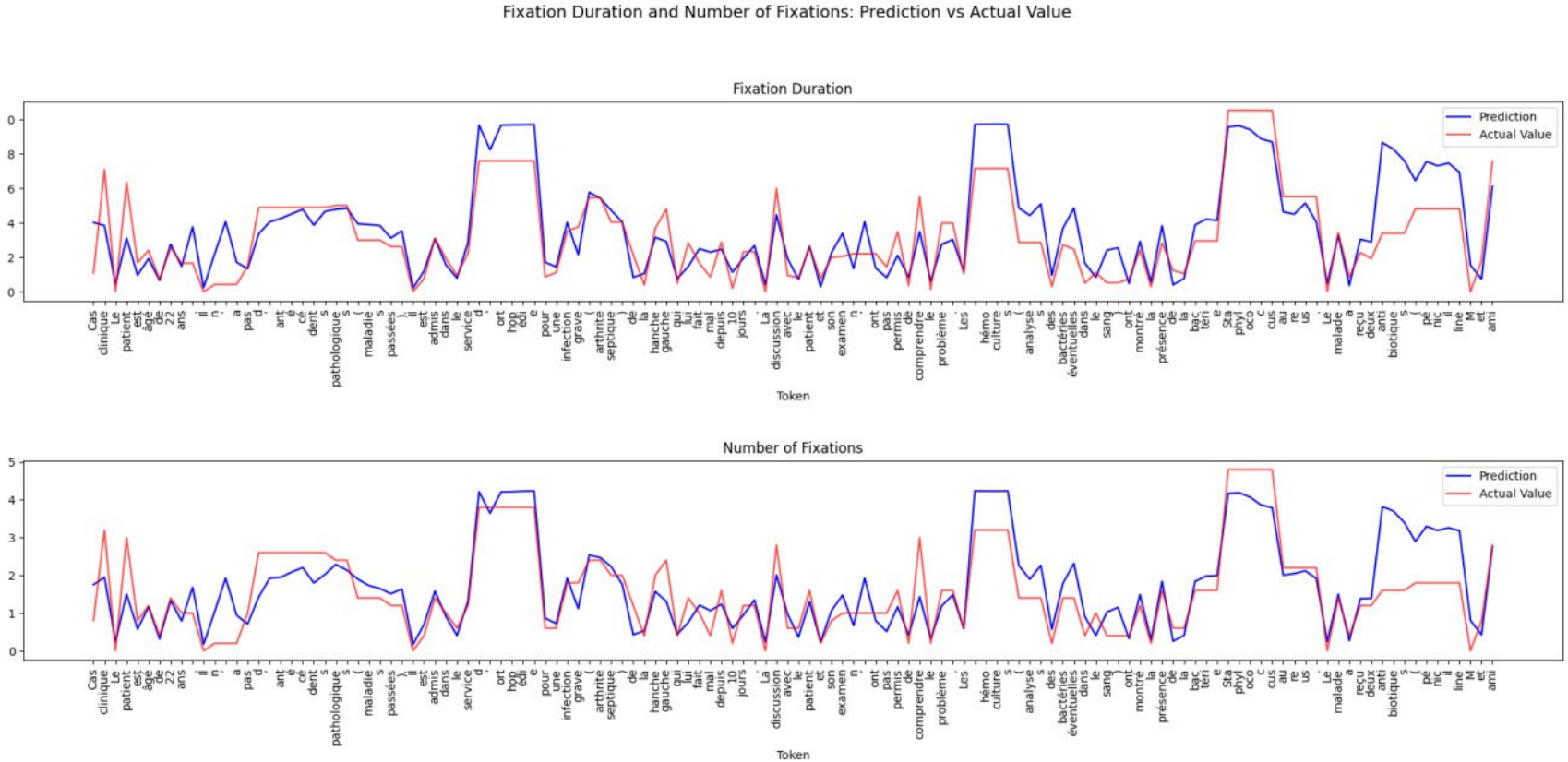
Metrics	MSE	RMSE	MAE	R^2 Score
Number of Fixations	0.0142	0.1191	0.0956	0.7896
Fixation Duration	0.4830	0.6950	0.5249	0.5732

General Text Original



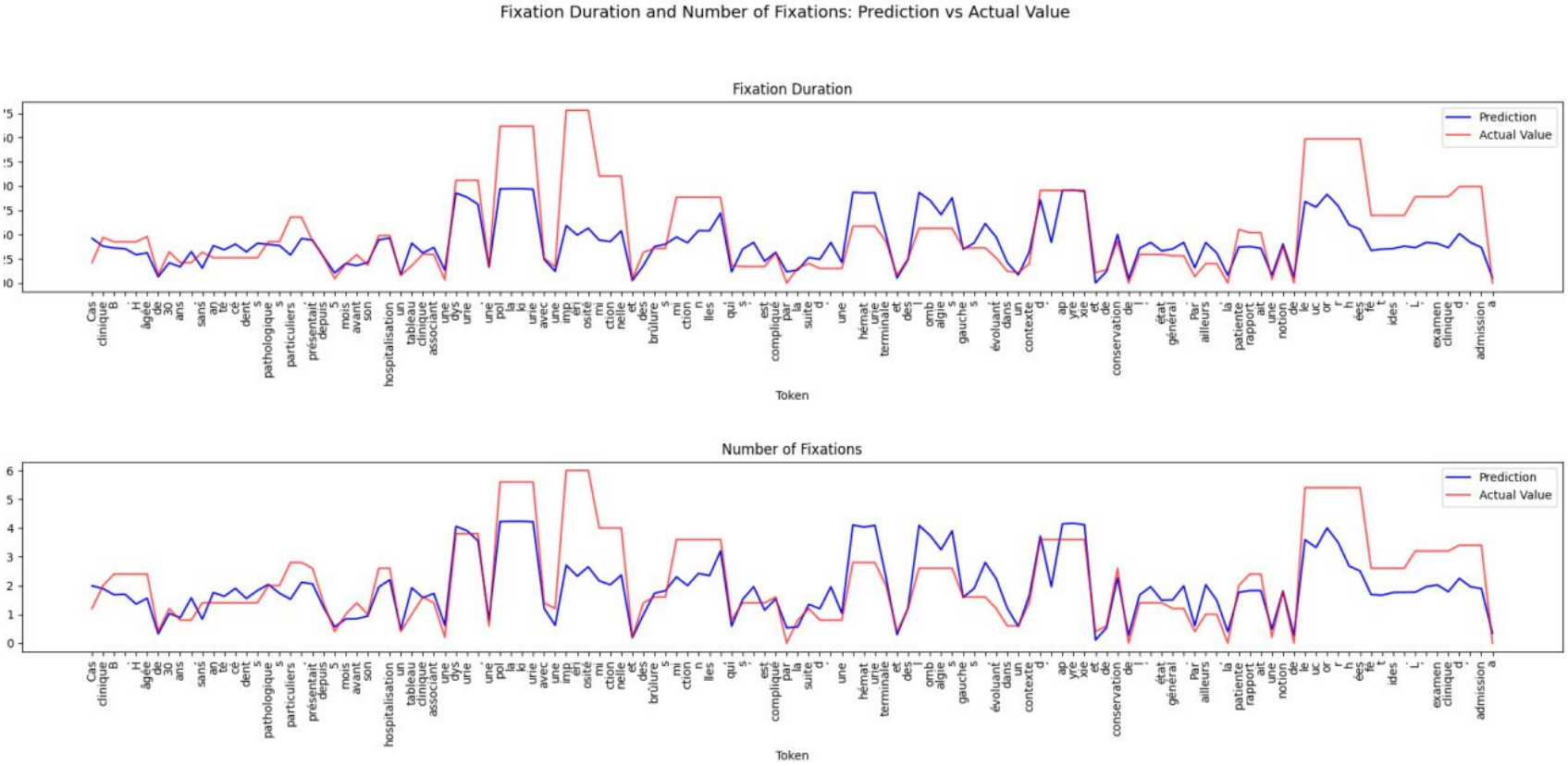
Metrics	MSE	RMSE	MAE	R^2 Score
Number of Fixations	0.0148	0.1218	0.0855	0.6203
Fixation Duration	0.3365	0.5801	0.4047	0.6213

Clinical Case Simplified



Metrics	MSE	RMSE	MAE	R^2 Score
Number of Fixations	0.0233	0.1528	0.1162	0.6538
Fixation Duration	0.4705	0.6859	0.5175	0.6583

Clinical Case Original



Metrics	MSE	RMSE	MAE	R^2 Score
Number of Fixations	0.1261	0.3550	0.2352	0.4116
Fixation Duration	1.0858	1.0420	0.7517	0.5504

Overview of Model Performance Metrics by Text Type

Text Type	Metric	MSE	RMSE	MAE	R^2 Score
Medical Orig.	Num. of Fixations	0.0142	0.1191	0.0956	0.7896
	Fix. Duration	0.4830	0.6950	0.5249	0.5732
Clin. Case Sim.	Num. of Fixations	0.0233	0.1528	0.1162	0.6538
	Fix. Duration	0.4705	0.6859	0.5175	0.6583
General Orig.	Num. of Fixations	0.0148	0.1218	0.0855	0.6203
	Fix. Duration	0.3365	0.5801	0.4047	0.6213
Clin. Case Orig.	Num. of Fixations	0.1261	0.3550	0.2352	0.4116
	Fix. Duration	1.0858	1.0420	0.7517	0.5504

Conclusion

- Conducting ongoing eye-tracking experiments
- Performing various statistical analyses on eye-tracking data
- Training models on two eye-tracking features:
  - duration of fixation
  - number of fixations
- Exploring CamemBERT for the predictions

- Complete eye-tracking data with more participants
- Training models on multiple eye-tracking features:
  - Extend to six fixation features (some are not correlated):
    - Time to first fixation
    - Duration of first fixation
    - Total duration of fixations
    - Average duration of fixations
    - Maximum duration of fixations
    - Number of fixations
  - Include additional eye-tracking features
- Develop models to predict the reader type:
  - explore other LLMs, such as LLAMA 3.1 and CamemBERT-bio, to improve eye-tracking prediction
  - distinguish between general and speech therapist students
- Define baselines

D. Aarsland, L. Batzu, G. Halliday, G. Geurtsen, C. Ballard, K. Ray Chaudhuri, and D. Weintraub. Parkinson disease-associated cognitive impairment. *Nature Reviews Disease Primers*, 7:47, 2021.

Neil K. Archibald, Sam B. Hutton, Michael P. Clarke, Urs P. Mosimann, and David J. Burn. Visual exploration in parkinson's disease and parkinson's disease dementia. *Brain*, 136(3):739–750, Mar 2013. doi: 10.1093/brain/awt005. URL `https://doi.org/10.1093/brain/awt005`.

M. Boucart, G. Bubbico, S. Szaffarczyk, and F. Pasquier. Animal spotting in alzheimer's disease: an eye tracking study of object categorization. *J Alzheimers Dis*, 39(1):181–189, 2014. doi: 10.3233/JAD-131331.

Anna Ekstrand, Mattias Nilsson, and Gustaf Öqvist Seimyr. Screening for reading difficulties: Comparing eye tracking outcomes to neuropsychological assessments. *Frontiers in Education*, 6, 03 2021. doi: 10.3389/feduc.2021.643232.

Matías García and Sandra Cano. *Eye Tracking to Evaluate the User eXperience (UX): Literature Review*, pages 134–145. 06 2022. ISBN 978-3-031-05060-2. doi: 10.1007/978-3-031-05061-9_10.

Aline Godfroid, Paula Winke, and Kathy Conklin. Exploring the depths of second language processing with eye tracking: An introduction. *Second Language Research*, 36:026765832092257, 06 2020. doi: 10.1177/0267658320922578.

Ooi K. and Khaidzir Ismail. The relationship between cognitive processing and reading. *Asian Social Science*, 7, 09 2011. doi: 10.5539/ass.v7n10p44.

Elaine Khoong, Gem Le, Mekhala Hoskote, Natalie Rivadeneira, Robert Hiatt, and Urmimala Sarkar. Health information-seeking behaviors and preferences of a diverse, multilingual urban cohort. *Medical care*, 57 Suppl 6 Suppl 2:S176–S183, 06 2019. doi: 10.1097/MLR.0000000000001050.

Young-Suk Grace Kim, Yaacov Petscher, and Christian Vorstius. Unpacking eye movements during oral and silent reading and their relations to reading proficiency in beginning readers. *Contemporary Educational Psychology*, 58:102–120, 2019.

Viktor Kuperman, Kimberly Matsuki, and Julie A. Van Dyke. Contributions of reader- and text-level characteristics to eye-movement patterns during passage reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(11): 1687–1713, Nov 2018. doi: 10.1037/xlm0000547. Epub 2018 Jul 19.

Francesca Magnabosco and Olaf Hauk. An eye on semantics: a study on the influence of concreteness and predictability on early fixation durations. *Language, Cognition and Neuroscience*, 39(3):302–316, 2023. doi: 10.1080/23273798.2023.2274558.

Alexandra Marquis, Meera Al Kaabi, Tommi Leung, and Fatima Boush. What the eyes hear: An eye-tracking study on phonological awareness in emirati arabic. *Frontiers in Psychology*, 11, 2020. ISSN 1664-1078. doi: 10.3389/fpsyg.2020.00452. URL `https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2020.00452`.

Damien C. Mézière, Lu Yu, Erik D. Reichle, Titus von der Malsburg, and Genevieve McArthur. Using eye-tracking measures to predict reading comprehension. *Reading Research Quarterly*, 58:425–449, 2023. doi: 10.1002/rrq.498.

Lubomira Novakova, Martin Gajdos, Jana Marková, Alice Martinkovicova, Zuzana Košutzká, Jana Svantnerova, Peter Valkovic, Zsolt Cséfalvay, and Irena Rektorova. Language impairment in parkinson's disease: fmri study of sentence reading comprehension. *Frontiers in Aging Neuroscience*, 15:1117473, 03 2023. doi: 10.3389/fnagi.2023.1117473.

Eszter Ronai, Yenan Sun, Alan Yu, and Ming Xiang. Integration of contextual-pragmatic and phonetic information in speech perception: An eye-tracking study. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 10, 11 2019. doi: 10.5334/labphon.186.

Lavinia Salicchi, Rong Xiang, and Yu-Yin Hsu. HkAmsters at CMCL 2022 shared task: Predicting eye-tracking data from a gradient boosting framework with linguistic features. In Emmanuele Chersoni, Nora Hollenstein, Cassandra Jacobs, Yohei Oseki, Laurent Prévot, and Enrico Santus, editors, *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, pages 114–120, Dublin, Ireland, May 2022. Association for Computational Linguistics. doi: 10.18653/v1/2022.cmcl-1.13. URL `https://aclanthology.org/2022.cmcl-1.13`.

Lavinia Salicchi, Emmanuele Chersoni, and Alessandro Lenci. A study on surprisal and semantic relatedness for eye-tracking data prediction. *Frontiers in Psychology*, 14, 2023. ISSN 1664-1078. doi: 10.3389/fpsyg.2023.1112365. URL `https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2023.1112365`.

Eva Soroli, Halima Sahraoui, and Carol Sacchett. Soroli, e., sahraoui h. & sacchett c. (2012). linguistic encoding of motion events in english and french: Typological constraints on second language acquisition and agrammatic aphasia. language, interaction & acquisition, 3(2): 261–287. doi: 10.1075/lia.3.2.05sor. *Language, Interaction and Acquisition*, 3:261–287, 12 2012. doi: 10.1075/lia.3.2.05sor.

K.M. Walle, J.E. Nordvik, F. Becker, T. Espeseth, M.H. Sneve, and B. Laeng. Unilateral neglect post stroke: Eye movement frequencies indicate directional hypokinesia while fixation distributions suggest compensational mechanism. *Brain Behav*, 9(1):e01170, Jan 2019. doi: 10.1002/brb3.1170. Epub 2018 Dec 12.