A Survey on Using Gaze Behaviour for Natural Language Processing

Sandeep Mathias[1], Diptesh Kanojia[1,2], Abhijit Mishra[3], and Pushpak Bhattacharyya[1]
[1]Indian Institute of Technology, Bombay
[2]IITB-Monash Research Academy
[3]Apple, Incorporated, Seattle, WA, USA

Outline

- Eye-Tracking Motivation & Terms
- Gaze Behaviour Corpora
- Motivation for Learning Gaze Behaviour
- Learning Gaze Behaviour for NLP Tasks
- Further Applications
- Concluding Remarks

Eye Tracking Motivation

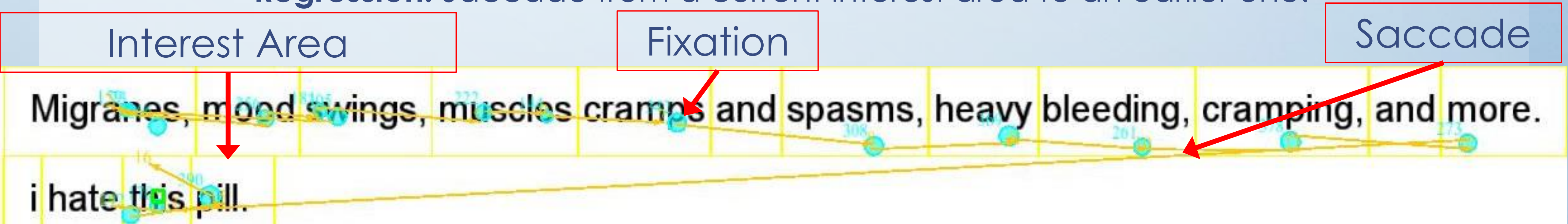
- Eye-tracking is a means of using cognitive information for solving different language processing and understanding tasks.
- Eye-tracking research is based on the Eye-Mind hypothesis:
  - There is no appreciable lag between what is fixated and what is processed.
    - Just & Carpenter. A Theory of Reading: From Eye Fixations to Comprehension. (1980)
- Gaze behaviour can be used to personalize the NLP system in solving tasks.

Eye-tracking Motivation

- Example: Sarcasm Understandability (Mishra et al. 2016)

Eye-tracking Terms

1. **Interest Area:** An interest area is the area of the screen which is of interest for us. Example: Words and the space around them.
2. **Fixation:** A fixation is an event where the eye is focused on a part of the screen.
3. **Saccade:** The movement of the eye from one fixation point to the next.
   - **Progression:** Saccade from a current interest area to a later one.
   - **Regression:** Saccade from a current interest area to an earlier one.



Gaze Behaviour Corpora: Languages

Dataset	Language	Stimulus	Subjects
Zang et al. (2018)	Chinese	90 sentences	35
Li et al. (2018)		15 documents	29
Cop et al. (2017)	Dutch	1 novel	33
Mak & Willems (2019)		3 stories	102
Kennedy et al. (2003)	French	20 documents	10
Nicenboim et al. (2016)	German	176 sentences	72
Kleigl et al. (2004)		144 sentences	55
Safavi et al. (2016)	Persian	136 sentences	40
Laurinavichuyte et al. (2017)	Russian	144 sentences	96
Nicenboim et al. (2016)	Spanish	212 sentences	79

Gaze Behaviour Corpora: Tasks

Dataset	Task	Stimulus	Subjects
Joshi et al. (2014)	Sentiment Analysis	1059 sentences	5
Mishra et al. (2016)	Sarcasm Understandability	1000 Tweets	7
Cheri et al. (2016)	Coreference Resolution	22 documents	14
Mishra et al. (2017)	Reading Complexity	32 documents	16
Mathias et al. (2018)	Text Quality Prediction	30 documents	20

Learning Gaze Behaviour: Motivation

- Recording gaze behaviour is costly in terms of time and money.
  - You must pay annotators.
  - You must supervise the annotators.
  - It takes time to calibrate and validate the eye-tracker for recording the gaze behaviour.
  - Noise must be cleaned up post recording the gaze data
  - … … …
- Solution:
  - Learning gaze behaviour

Learning Gaze Behaviour: Solutions

- **Type Aggregation:** For a given token (T), the value of the corresponding gaze behaviour feature's value (F) is the mean value of that feature value for the token across the corpus.

- **Multi-Task Learning:** Learning gaze behaviour features are the auxiliary tasks while solving the given NLP problem is the primary task.

Normalizing Gaze Behaviour

- Readers read at different speed. So, gaze data should be normalized.

- **Min-Max Normalization:** For a ***given reader***, normalize the feature values of each feature to the range of [0,1].

- **Binning:** For a ***given reader***, assign the feature value of a given gaze feature to a given bin.

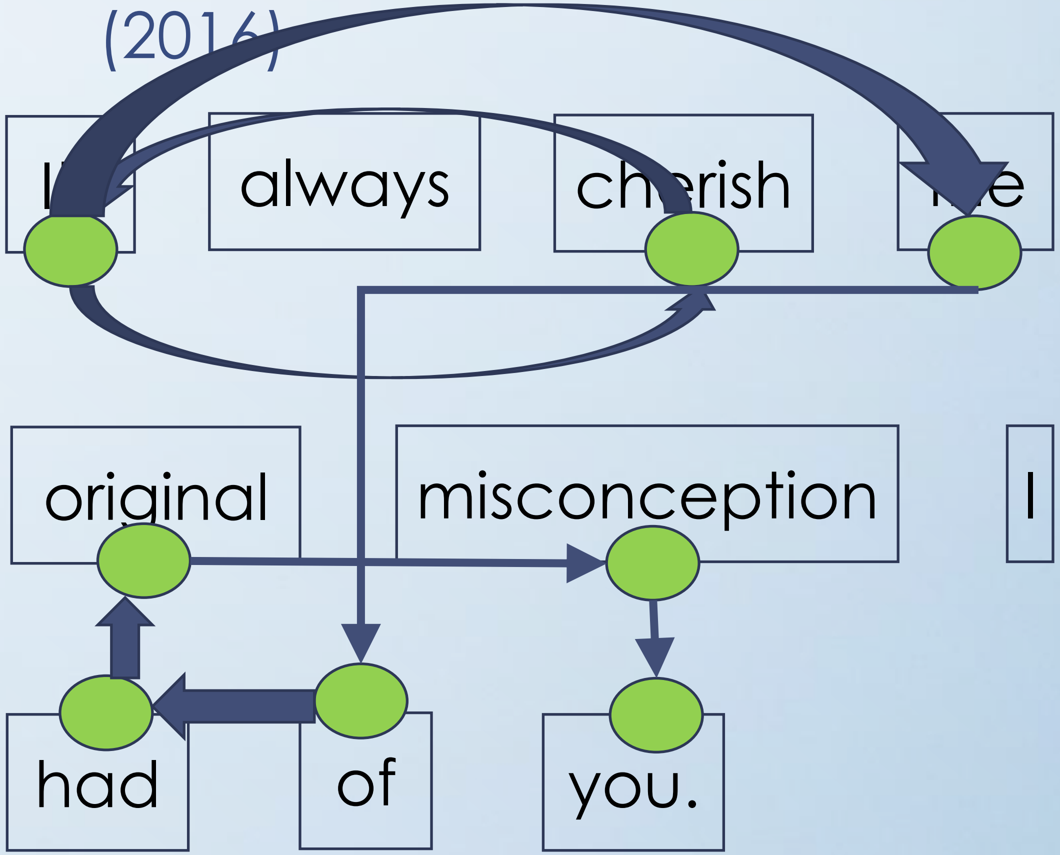
Learning Gaze Behaviour for NLP Tasks

- Predicting Fixations While Reading
- Predicting Grammatical Functions
- Text Simplification
- Part-of-Speech Tagging
- Readability
- Sentiment Analysis
- Sequence Classification
- Named Entity Recognition

Predicting Fixations While Reading

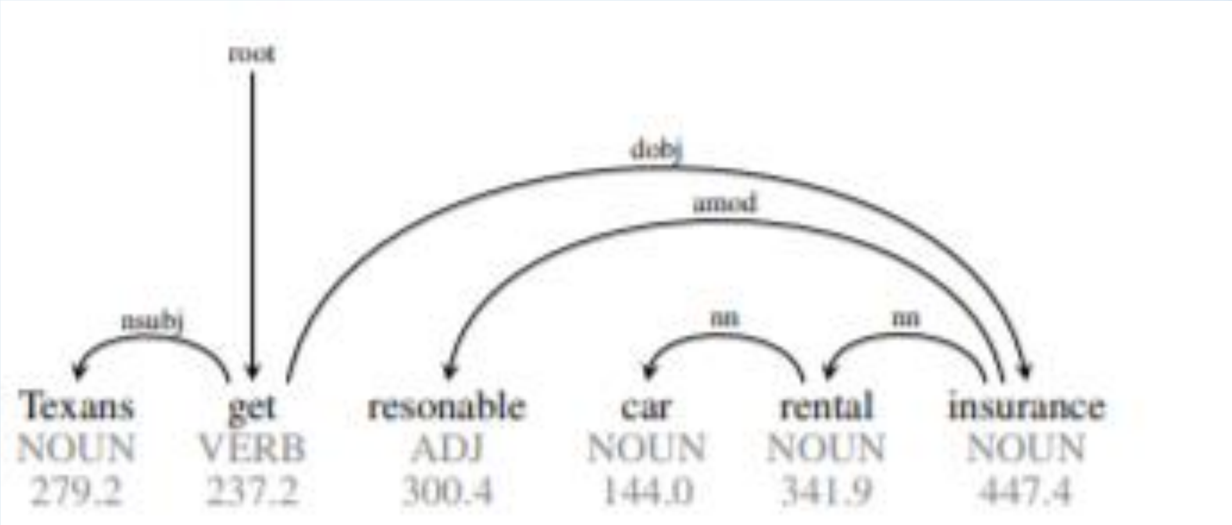
- Nilsson & Nivre (2009) used a transition-based approach to predict the next fixation.
  - Used features like token length, token frequency, next token length, etc.
- Matthies & Sogaard (2013) use a linear CRF model to predict the next fixation.
  - Used features like word length & word probability.

- Example source: Mishra et al. (2016)



Predicting Grammatical Functions

- Barrett & Sogaard (2015) use gaze features to predict grammatical functions.
  - Gaze features are learnt using type aggregation from the Dundee Corpus (Kennedy et al. (2003)).
- Figure shows the dependency parse of the sentence with mean fixation durations per word.
  - Barrett & Sogaard (2015)



Text Simplification

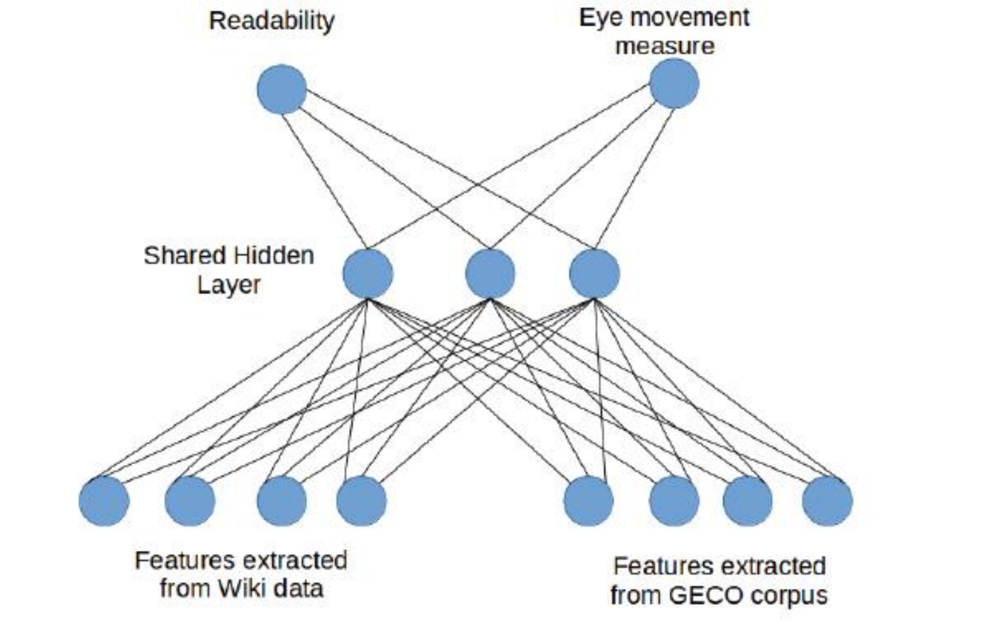
- Klerke et al. (2016) use gaze behaviour to learn to simplify text by compressing sentences.
  - They used multi-task learning, learning the first fixation duration and regression durations as auxiliary tasks and compressing the sentence as the primary task using Bi-LSTMs.
- Example:
  - Input: Intel would be building car batteries, expanding its business model beyond its core strength, the company said in a statement.
  - Output: Intel would be expanding its business model beyond its core strength.

Part-of-Speech Tagging

- Barrett et al. (2016a) and Barrett et al. (2016b) describe an approach to solve PoS tagging using type aggregations from the Dundee Corpus in a monolingual (Barrett et al. (2016a)) and cross-lingual (Barrett et al. (2016b)) setting using a Hidden Markov Model.
- They used a variety of features such as
  - Basic gaze features (Eg. Dwell Time)
  - Early gaze features (Eg. First Fixation Duration)
  - Late gaze features (Eg. Regression-to Duration)
  - Context gaze features (Eg. Fixation probability of nearby words)
  - Text features (Eg. Word length, corpus frequency)

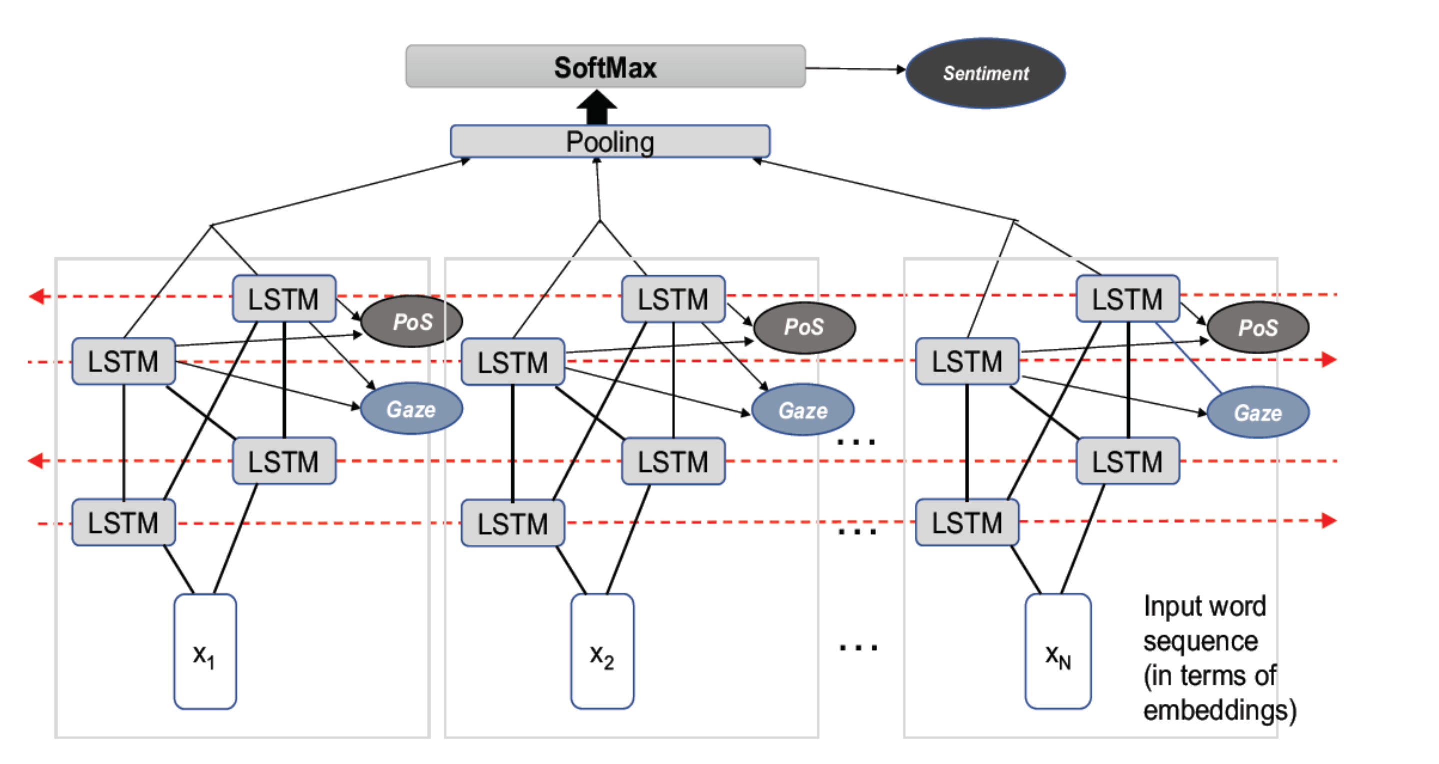
Readability

- Gonzalez-Garduno and Sogaard (2018) used a multi-task learning multi-layer perceptron with gaze behaviour learnt as an auxiliary task to improve prediction of readability.
- Learnt features from the Dundee Corpus and the Ghent Eye-tracking Corpus (GECO) (Cop et al. (2017).
  - Source:Gonzalez-Garduno & Sogaard (2018)



Sentiment Analysis

- Mishra et al. (2018) perform sentiment analysis by learning Part-of-Speech tagging and gaze behaviour at run time as auxiliary tasks.
  - Source: Mishra et al. (2018)

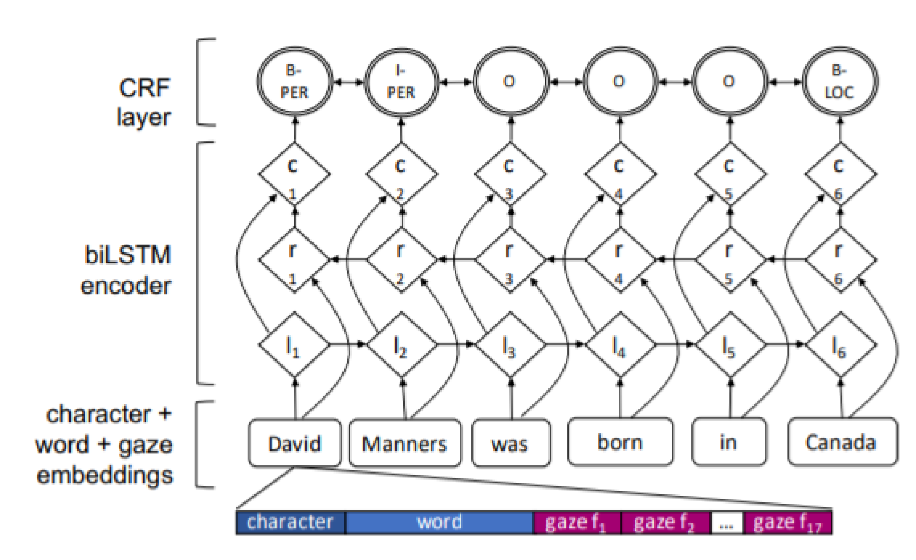


Sequence Classification

- Barrett et al. (2018) used a multi-task learning approach to solve sentiment analysis, grammatical error detection and hate speech detection.
- They used an attention-based system where the attention learnt was in the form of gaze behaviour as an auxiliary task.
- Their system learnt the gaze behaviour at the token level and the task label at the sentence level.

Named-Entity Recognition

- Hollenstein & Zhang (2019) use a BiLSTM encoder and CRF layer. They use word embeddings, concatenated by character embeddings and 17 gaze features, namely Basic, Early, Late & Context features.
- The features are type aggregated from the Dundee, ZuCo (Hollenstein et al. (2018)) and GECO (Cop et al. (2017)).
  - Source: Hollenstein & Zhang (2019)



Further Proposed Applications

- 1. Complex Word Identification (CWI)
- 2. Automatic Essay Grading (AEG)
  - Currently there is a publication of learning gaze behaviour for AEG (Mathias et al. (2020)).

Complex Word Identification

- CWI is identifying whether a word / phrase is complex in the given context.
- It is important for the task of lexical simplification.
- Gaze behaviour research has been done on quantifying complexity:
  - Translation Complexity – Mishra et al. (2013)
  - Sentiment Annotation Complexity – Joshi et al. (2014)
  - Scanpath Complexity – Mishra et al. (2017)
- But no work has been done on ***learning gaze behaviour*** for complex word identification.

Automatic Essay Grading

- Scoring a text written in response to a topic, called the essay prompt.
- Mathias et al. (2018) showed that gaze behaviour can help in ***predicting the quality rating of a text*** given by a reader.
- Mathias et al. (2020) is a recent work which shows a solution to automatic essay grading where gaze behaviour is learnt as an auxiliary task.

Concluding Remarks

- Gaze behaviour has been shown to aid in solving multiple NLP tasks (Mishra & Bhattacharyya (2018)).
- However, collecting gaze behaviour at run time is not feasible.
- In order to use gaze behaviour, we utilize different approaches, like multi-task learning, using type aggregated values, etc.
- Gaze behaviour has been learnt for solving multiple NLP tasks such as PoS tagging, sentence compression, Named-Entity Recognition, sentiment analysis, automatic essay grading, etc.

Thank You!

Questions?

References (1 / 3)

- Maria Barrett & Anders Sogaard. 2015. Using Reading Behavior to Predict Grammatical Functions. In *Proceedings of the 6th Workshop on Cognitive Aspects of Computational Language Learning*, pages 1–5.
- Maria Barrett, Joachim Bingel, Frank Keller, and Anders Sogaard. 2016a. Weakly Supervised Part-of-Speech Tagging Using Eye-tracking Data. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 579–584.
- Maria Barrett, Frank Bingel, and Anders Sogaard. 2016b. Cross-lingual Transfer of Correlations Between Parts of Speech and Gaze Features. In *Proceedings of the 26th International Conference on Computational Linguistics*, pages 1330–1339.
- Maria Barrett, Joachim Bingel, Nora Hollenstein, Marek Rei, and Anders Sogaard. 2018. Sequence Classification with Human Attention. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pages 302–312.
- Joe Cheri, Abhijit Mishra, and Pushpak Bhattacharyya. 2016. Leveraging Annotators' Gaze Behaviour for Coreference Resolution. In *Proceedings of the 7th Workshop on Cognitive Aspects of Computational Language Learning*, pages 22–26.
- Uschi Cop, Nicolas Dirix, Denis Drieghe, and Wouter Duyck. 2017. Presenting GECO: An Eye-tracking Corpus of Monolingual and Bilingual Sentence Reading. *Behavior Research Methods*, 49(2):602–615.
- Ana Gonzalez-Garduno and Anders Sogaard. 2018. Learning to Predict Readability Using Eye-movement Data from Natives and Learners. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, pages 5118 – 5124.
- Nora Hollenstein and Ce Zhang. 2019. Entity Recognition at First Sight: Improving NER with Eye Movement Information. In *Proceedings of the 14th Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 1–10. ACL, 2019.
- Nora Hollenstein, Jonathan Rotsztejn, Marius Troendle, Andreas Pedroni, Ce Zhang, and Nicolas Langer. 2018. ZuCo, a Simultaneous EEG and Eye-tracking Resource for Natural Sentence Reading. *Scientific Data*, 5(1):1–13.
- Aditya Joshi, Abhijit Mishra, Nivvedan Senthamilselvan, and Pushpak Bhattacharyya. 2014. Measuring Sentiment Annotation Complexity of Text. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 36–41.

References (2 / 3)

- Alan Kennedy, Robin Hill, and Joel Pynte. 2003. The Dundee Corpus. In *Proceedings of the 12th European Conference on Eye Movement*.
- Sigrid Klerke, Yoav Goldberg, and Anders Sogaard. 2016. Improving Sentence Compression by Learning to Predict Gaze. In *Proceedings of the 12th Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1528–1533.
- Reinhold Kliegl, Ellen Grabner, Martin Rolfs, and Ralf Engbert. 2004. Length, Frequency, and Predictability Effects of Words on Eye Movements in Reading. *European Journal of Cognitive Psychology*, 16(1-2):262– 284.
- Anna Laurinavichyute, Irina Sekerina, Svetlana Alexeeva, and Kristine Bagdasaryan. 2017. Russian Sentence Corpus: Benchmark Measures of Eye Movements in Reading in Cyrillic.
- Xiangsheng Li, Yiqun Liu, Jiaxin Mao, Zexue He, Min Zhang, and Shaoping Ma. 2018. Understanding Reading Attention Distribution During Relevance Judgement. In *Proceedings of the 27th ACM Conference on Information and Knowledge Management*, pages 733–742.
- Marloes Mak & Roel Willems. 2019. Mental Simulation During Literary Reading: Individual Differences Revealed with Eye-tracking. *Language, Cognition and Neuroscience*, 34(4):511–535.
- Marcel Just & Patricia Carpenter. 1980. A Theory of Reading: From Eye Fixations to Comprehension. *Psychological Review*, 87(4):329.
- Sandeep Mathias, Diptesh Kanojia, Kevin Patel, Samarth Agrawal, Abhijit Mishra, and Pushpak Bhattacharyya. 2018. Eyes are the Windows to the Soul: Predicting the Rating of Text Quality Using Gaze Behaviour. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2352–2362.
- Sandeep Mathias, Rudra Murthy, Diptesh Kanojia, and Pushpak Bhattacharyya. 2020. Happy are Those who Grade Without Seeing: A Multi-Task Learning Approach to Grade Essays Using Gaze Behaviour. In *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference in Natural Language Processing*, pages 858 – 872.

References (3 / 3)

- Franz Matthies & Anders Sogaard. 2013. With Blinkers on: Robust Prediction of Eye Movements Across Readers. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 803–807.
- Abhijit Mishra and Pushpak Bhattacharyya. 2018. *Cognitively Inspired Natural Language Processing: An Investigation Based on Eye-Tracking*. Springer.
- Abhijit Mishra, Pushpak Bhattacharyya, and Michael Carl. 2013. Automatically Predicting Sentence Translation Difficulty. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 346–351.
- Abhijit Mishra, Diptesh Kanojia, and Pushpak Bhattacharyya. 2016. Predicting Readers' Sarcasm Understandability by Modeling Gaze Behaviour. In *Proceedings of the 30th AAAI Conference on Artificial Intelligence*, pages 3747–3753.
- Abhijit Mishra, Diptesh Kanojia, Seema Nagar, Kuntal Dey, and Pushpak Bhattacharyya. 2017. Scanpath Complexity: Modeling Reading Effort Using Gaze Information. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 4429–4436.
- Abhijit Mishra, Srikanth Tamilselvam, Riddhiman Dasgupta, Seema Nagar, and Kuntal Dey. 2018. Cognition-Cognizant Sentiment Analysis with Multitask Subjectivity Summarization Based on Annotators' Gaze Behavior. In *Proceedings of the 32nd AAAI Conference on Artificial Intelligence*, pages 5884 – 5891.
- Bruno Nicenboim, Pavel Logacev, Carolina Gattei, and Shravan Vasishth. 2016. When High-Capacity Readers Slow Down and Low-Capacity Readers Speed Up: Working Memory and Locality Effects. *Frontiers in Psychology*, 7:280.
- Mattias Nilsson & Joakim Nivre. Learning Where to Look: Modeling Eye Movements in Reading. 2009. In *Proceedings of the 13th Conference on Computational Natural Language Learning*, pages 93–101.
- Molood S Safavi, Samar Husain, and Shravan Vasishth. 2016. Dependency Resolution Difficulty Increases with Distance in Persian Separable Complex Predicates: Evidence for Expectation and Memory-based Accounts. *Frontiers in Psychology*, 7:403.
- Chuanli Zang, Ying Fu, Xuejun Bai, Guoli Yan, and Simon P Liversedge. Investigating Word Length Effects in Chinese Reading. 2018. *Journal of Experimental Psychology: Human Perception and Performance*, 44(12):1831–1841.