

Introduction

- Most eye movement studies in reading examine content words (**CW**).
- Function words (**FW**) such as “the” are shorter, carry less semantic information, are more frequent compared to **CW**, and are also skipped more often (Rayner, 1998).
- The article “the” is often skipped even when used as a grammatically incorrect parafoveal preview for the target word (Angele & Rayner, 2013), indicating differential processing for **FW** compared to **CW**.
- In some languages, such as Brazilian Portuguese (BP), **FW** can carry more semantic information (compared to English) in the form of gender and number marks.
- Do observations on FW that carry little information (e.g. the) replicate in languages where FW carry more information?**

Results and Discussion

Table 1. Means and standard deviations for eye movement measures for content and function words.

Measure	CW		FW	
	Mean	SD	Mean	SD
Skipping (%)	0.21	0.4	0.61	0.48
FFD (ms)	214	68	202	69
GD (ms)	258	111	226	108
GPT (ms)	377	330	316	301

We found main effects of predictability and word length on all measures in the expected direction.

No effects of word class on Gaze Duration and Go Past Time when the other predictors were included in the models.

For **skipping rates**, we found three interactions.

- Predictability and word length** - shorter words showed a stronger effect of predictability, while there was barely any effect of predictability on longer words (**Fig. 1**). There were floor effects on long words, which were rarely if ever skipped.
- Predictability and word class** showed a stronger effect of predictability on **CW** (**Fig. 2**). We believe that predicting a word's part of speech may be enough to influence the skipping of **FW**, which would mean a full-word prediction isn't necessary. Luke and Christianson (2016) showed that partial predictability was enough to facilitate skipping of **FW**, but not **CW**.
- Word length and word class**, where shorter **FW** were skipped more often than same-sized **CW**, while the opposite was also somewhat true for longer words (**Fig. 3**). When words get longer, it becomes harder to identify its class in the parafovea and skipping rates become more similar between **FW** and **CW**.

For **FFD**, there was one three-way interaction of **predictability, length, and word class** where very predictable **FW** words were processed very quickly, regardless of length, while length had a stable effect on **CW** regardless of length (**Fig. 4**).

We believe these results on **FFD** might be consequence of the high word skipping where some of the instances when the eye do land on the word are due to mislocated fixations (Nuthmann, Engbert & Kliegl, 2005). When the eyes land on a highly predictable **FW** intended to be skipped, the eyes quickly continue to the next word. Later measures would not be sensitive to this phenomenon.

Example of FW in BP

In BP, the words “algum/alguns/alguma/algumas”, “some” in English, can be masculine and feminine, and singular and plural.

[...] era ponto de reunião para **alguns** rapazes [...] – [...] (it) was a place of reunion for **some** men [...]. Masculine and plural.

[...] e **algumas** moças elegantes [...] – [...] and **some** fancy ladies [...]. Feminine and plural.

Methods and Material

We analysed data from the RASTROS Corpus of natural reading in BP (Vieira, 2020).

- Cloze data for every word, except the first word from 50 paragraphs - 393 participants.
- Forty-nine participants read all paragraphs – EyeLink 1000 (SR Research).
- 2494 words (978 function words).
- 49 words on average per paragraph.
- 4.96 average word length.

Objective

- To analyse the effects of word predictability, word class (**FW** versus **CW**), and length in Brazilian Portuguese.

Analyses

- A high negative correlation between length and frequency (-0.67) made it problematic to include both as predictors in a model.
- This correlation is not unexpected as the paragraphs were not manipulated.
- All **FW** had high or very high frequency. There is evidence showing that there is little to no frequency effect between high and very high-frequency words (Brysbaert, Mander & Keuleers, 2018).
- We therefore only used **CW** that had equivalent frequency as the **FW** (removing 36% of all **CW** from analyses). This allowed us to remove frequency as a predictor.
- Word length was centred around the average word length (4.96).

Figures

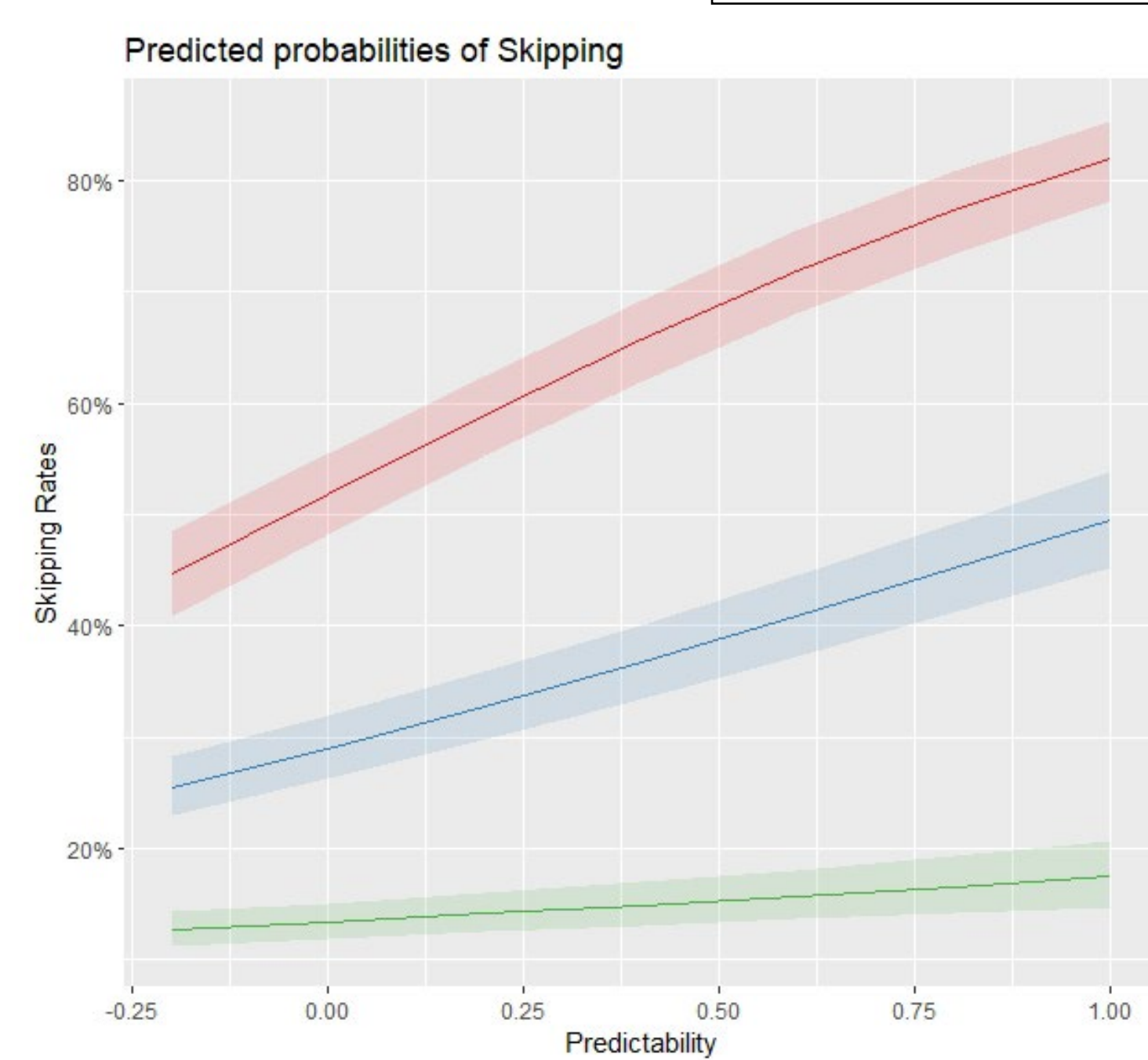


Figure 1. Word predictability and word length on skipping rates. The X-axis has predictability levels and the Y-axis has skipping rates. Colours indicate word length from short (red) to long (green).

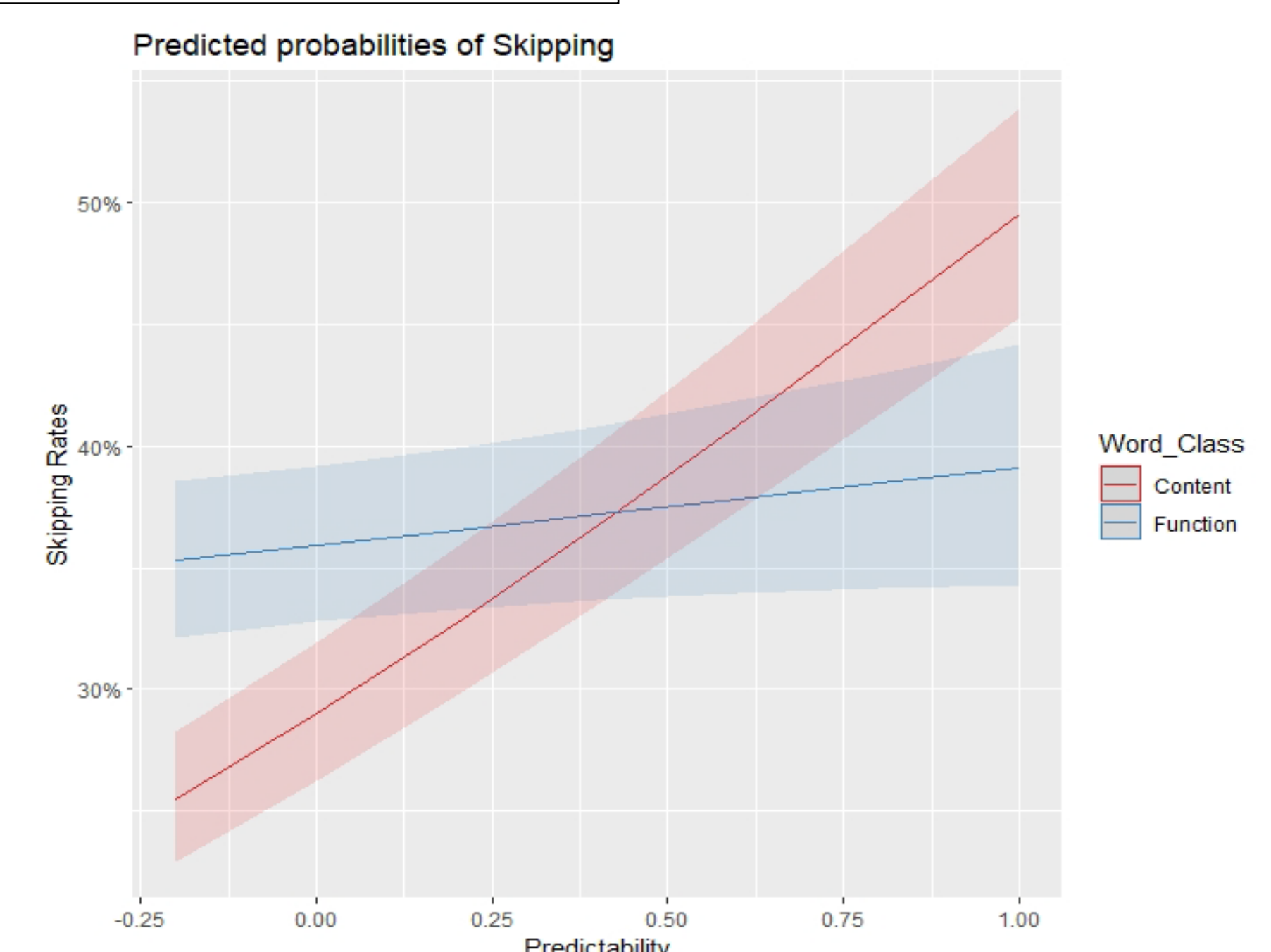


Figure 2. Word predictability and word class on skipping rates. The X-axis has predictability levels and the Y-axis has skipping rates. The blue line represents FW and the red CW.

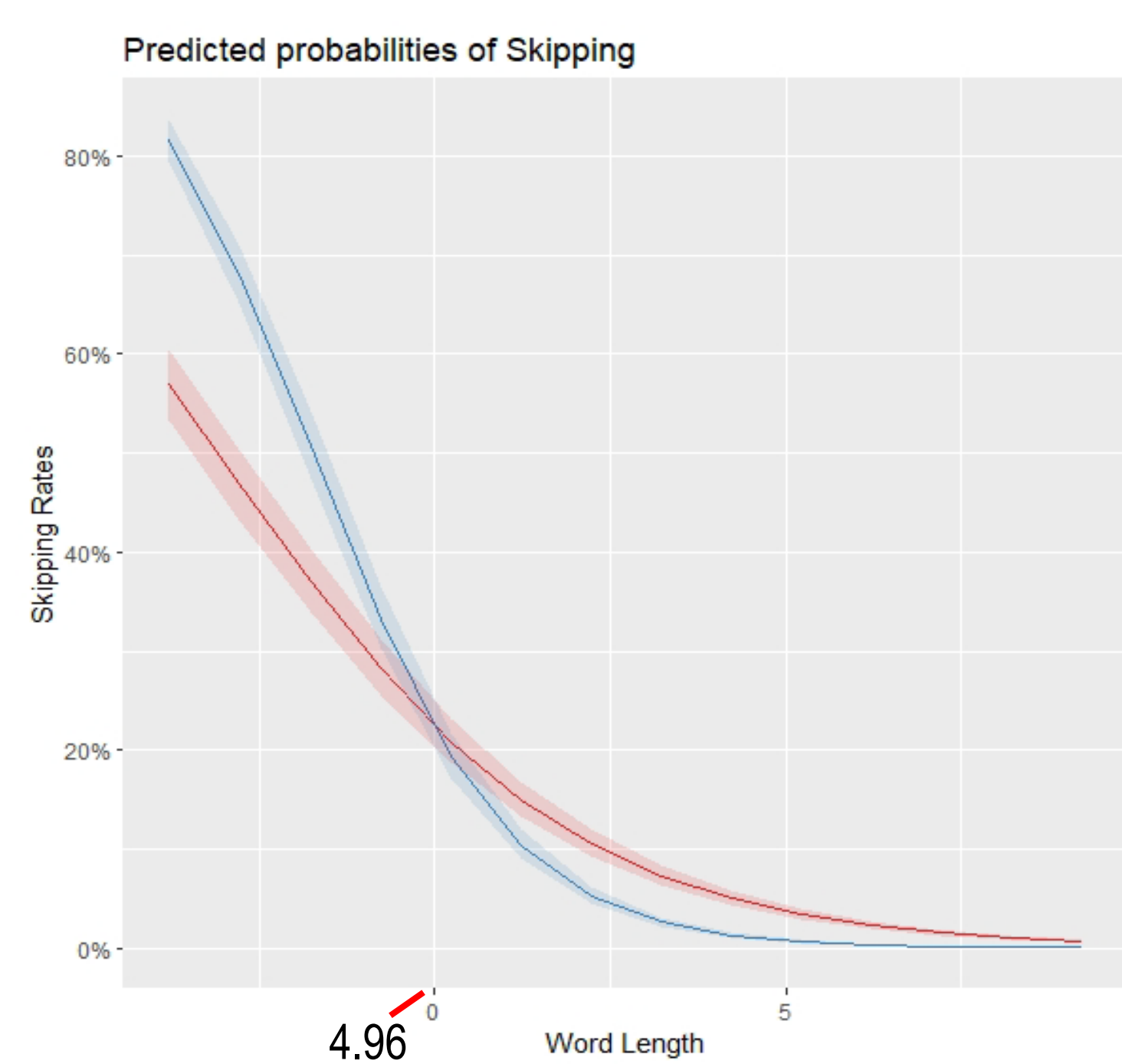


Figure 3. Word length and word class on skipping rates. The X-axis has word length, centred around the average length of 4.96, and the Y-axis has skipping rates. The blue line represents FW and the red CW.

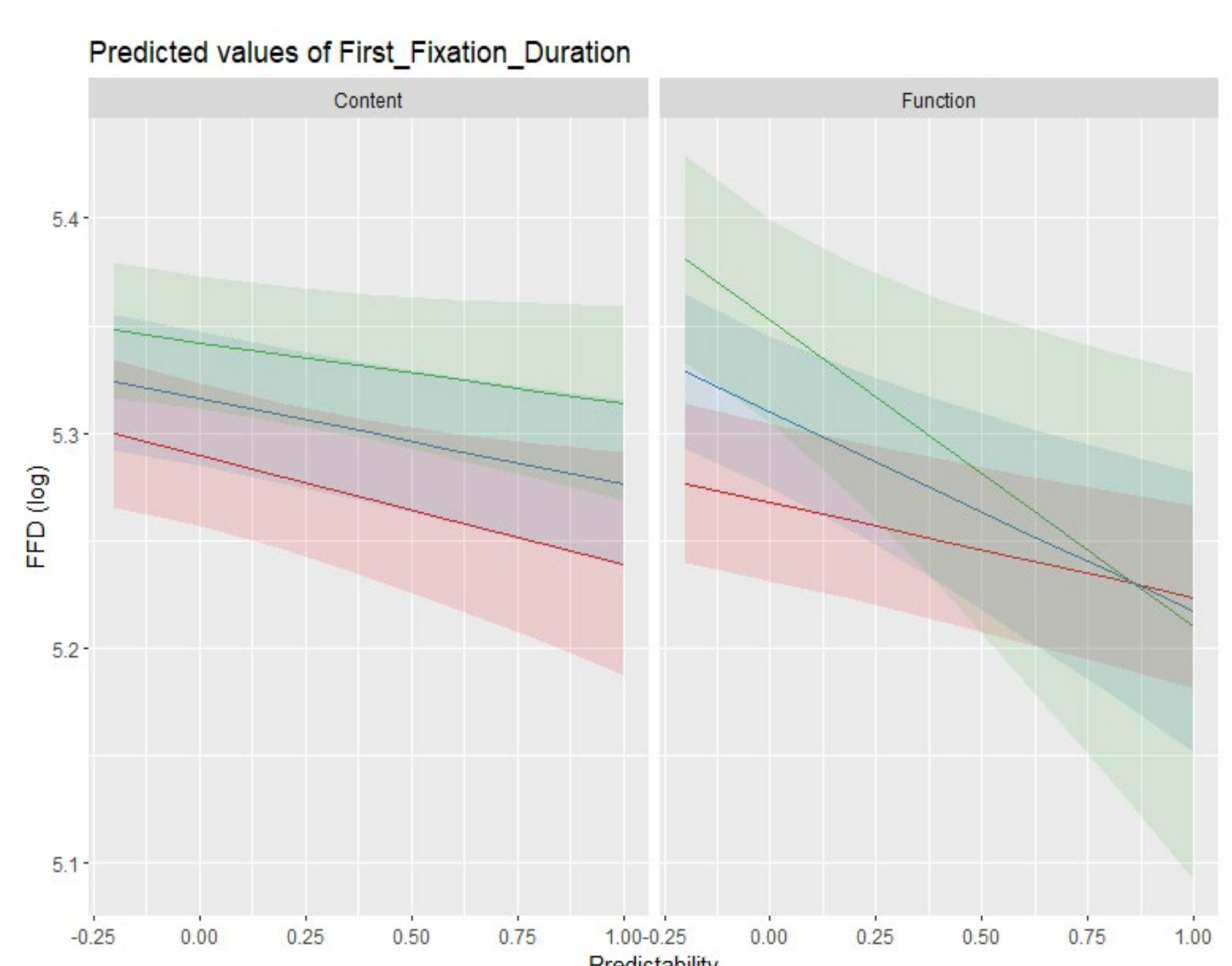


Figure 4. Word predictability, word length, and word class on FFD. The X-axis has predictability levels and the Y-axis has FFD (log). The colours indicate word length from short (red) to long (green), centred around the average of 4.96. The box on the left shows the results for CW and the box on the right for FW.

Conclusion

- There were differences between classes but only on very early processing, more specifically, shorter **FW** were skipped more often.
- We found limited differences in FFD, probably due to high skipping rates accompanied by mislocated fixations.
- Word class didn't predict any of the later eye movements.
- The heavier semantic load that **FW** have in BP is probably responsible for these results.
- Future research focusing on orthogonally manipulated frequency and length of **CW** and **FW** could contribute to this discussion.
- Skipping is assumed to be based on partial processing in the parafovea. The decision to skip seems to be influenced by word class even though processing time, as indicated by fixation times, which is assumed to reflect identification, is similar between **CW** and **FW**. This indicates word skipping and fixation times reflect differential stages in word processing (e.g. Drieghe, 2008).
- These results are compatible with Schmauder, Morris and Poynor (2000), who observed little to no differences in processing costs between word classes when **FW** and **CW** are matched in length and frequency. Although this latter study would need to be replicated (stimuli are not provided).

References

- Angele, B., & Rayner, K. (2013). Processing in the parafovea: Are articles skipped automatically? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(2), 649–662. <https://doi.org/10.1037/a0029294>
- Brysbaert, M., Mander, P., & Keuleers, E. (2018). The Word Frequency Effect in Word Processing: An Updated Review. *Current Directions in Psychological Science*, 27(1), 45–50. <https://doi.org/10.1177/0963721417727521>
- Drieghe, D. (2008). Foveal processing and word skipping during reading. *Psychonomic Bulletin & Review*, 15(4), 856–860. <https://doi.org/10.3758/PBR.15.4.856>
- Luke, S. G., & Christianson, K. (2016). Limits on lexical prediction during reading. *Cognitive Psychology*, 88, 22–60. <https://doi.org/10.1016/j.cogpsych.2016.06.002>
- Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated during reading and the inverted optimal viewing position effect. *Vision Research*, 45(fixations17), 2201–2217. <https://doi.org/10.1016/j.visres.2005.02.014>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Staub, A., Dodge, S., & Cohen, A. L. (2018). Failure to detect function word repetitions and omissions in reading: Are eye movements to blame? *Psychonomic Bulletin & Review*, 26(2), 340–346. <https://doi.org/10.3758/s13423-018-1492-z>
- Schmauder, A. R., Morris, R. K., & Poynor, D. V. (2000). Lexical processing and text integration of function and content words: Evidence from priming and eye fixations. *Memory & Cognition*, 28(7), 1098–1108. <https://doi.org/10.3758/BF03211811>
- Vieira, J. M. M. (2020). The Brazilian Portuguese eye tracking corpus with a predictability study focusing on lexical and partial prediction. Master's thesis, Federal University of Ceará (UFC), Universidade Federal do Ceará, Biblioteca Universitária, <http://www.repositorio.ufc.br/handle/riufc/55798>