An eye-tracking database of natural reading in Italian children

Valentina N. Pescuma[1], Maria Ktori[1], Benedetta Cevoli[1,2], Eleonora Lomi[1,3], Francesca Franzon[1], Davide Crepaldi[1]

[1] International School for Advanced Studies (SISSA), Trieste, Italy. [2] Royal Holloway University of London, UK. [3] University of Oxford, UK.

✉ vpescuma@sissa.it

Get PDF here

Background

- Reading proficiency may build up through the chunking of lower-level units (e.g., letters) into larger ones (e.g., words and morphemes)[1,2].
- Morpho-orthographic chunking in adults may be interpreted similarly — morphology drives regularities in letter co-occurrence within words[3,4,5], which the reading system may exploit to facilitate visual word identification.
- In this perspective, reading may be conceived as a form of **statistical learning**.

Aims and Hypotheses

- To identify **statistical learning proxies** in **developing readers of Italian** (3rd – 6th graders).
- Focus on **nGram** frequency. Other possible indexes (e.g., transitional probabilities, word predictability) are currently under investigation.
- Age expected to play a role in the development of sensitivity to statistical learning cues in reading.
- Data to be made available as one of the first **eye-tracking databases** in children.

Methods

Participants:
- 112 (63 M) native Italian speakers; age range: 8–12 years (mean=9.85, SD=1.13).

Procedure:
- Natural reading task on texts from kids' books.
- Simple 2-AFC comprehension questions after every other excerpt displayed.
- Eye movements recorded through a tower-mount Eyelink 1000 Plus eye-tracker.
- Computerized cloze probability task, currently under analysis.

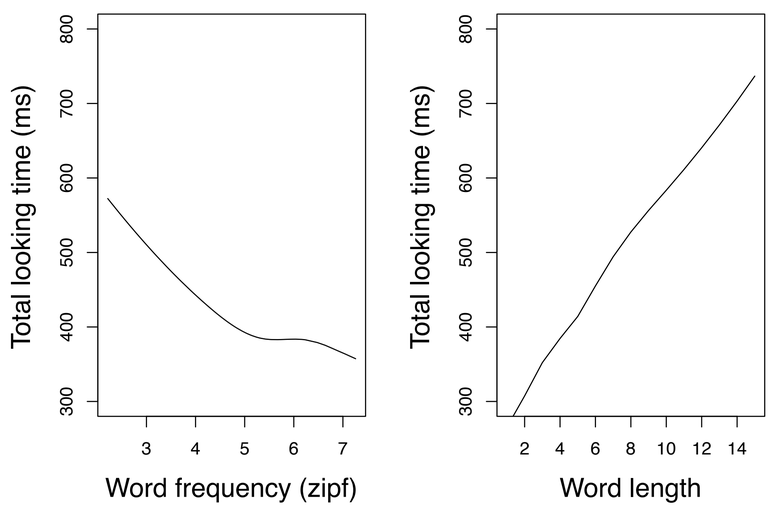
Additional assessment:
- Reading proficiency test (MT test – Speed and Accuracy[6]).
- Non-verbal intelligence test (Raven CPM-47[7]).

Stimuli features

1546 tokens
749 different words
609 different lemmas
12 parts of speech

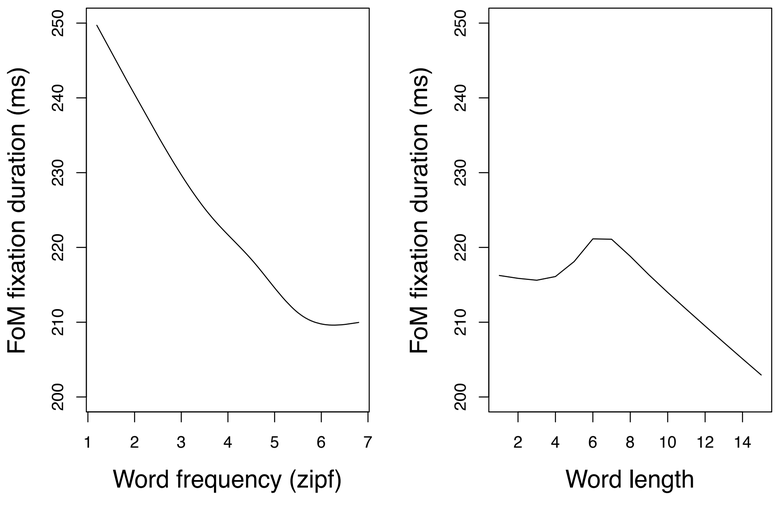
Results

Total Looking Time (TLT)



In line with previous data[8,9], benchmark effects on total looking of **word length** (F[4, 59593] = 370.75, p<.001) and **word frequency** (F[4, 59593] = 105.12, p<.001).

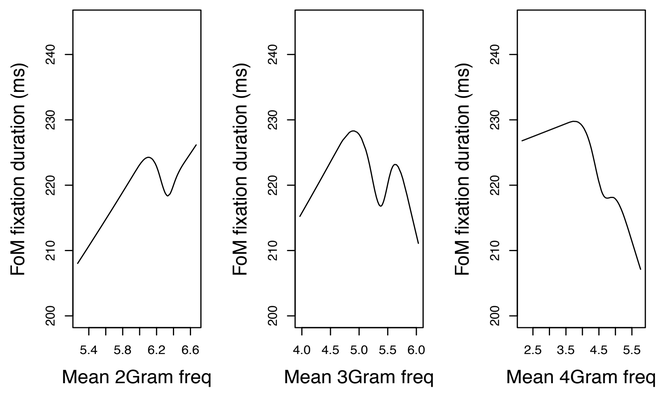
First-of-Many Fixation Duration (FoMFD)



Effects of **word frequency** (F[4, 19012]=14.11, p<.001) on FoMFD. A trend for **word length** (F[4, 19012]=1.82, p=.121).

Mean nGram Frequency

- **Size gradient**: no effect of 2Gram frequency (F[4, 16410]=0.85, p=.493); **3Grams** slightly better (F[4, 16410]=2.36, p=.051); significant effect of **4Grams** (F[4, 16410]=4.93, p=.001). Figures on the right.
- Size gradient particularly strong with reference to **early processing measures** (FoMFD).



- **4Gram** effects on FoMFD tend to **fade** when **word frequency** is also considered (F[4,16410]=1.39, p=.134).
- nGram effects on FoM **not modulated by age**, particularly for larger nGrams (e.g., 4Grams, p=.436). Figure on the right.

FoMFD

Conclusions

- Sensitivity to letter co-occurrence statistics from a very young age.
- Higher sensitivity to larger clusters than to small ones.
- High grapheme–phoneme correspondence in Italian — does phonology contribute to the statistical regularities, that are "inherited" by orthography once the mapping is learnt? Cross-linguistic research is needed.
- What is coded is still not clear. Lexical mediation in the present analysis (interest areas are words) but this shouldn't be taken for granted.

1, Castles & **Nation** (2006), Hove, UK: Psychology Press, 151–179. **2, Grainger** & **Ziegler** (2011), Frontiers in Psychology, 2(54):1–13. **3, Longtin** et al. (2003), Language and Cognitive Processes, 18:313–334. **4, Rastle** et al. (2004), Psychonomic Bulletin and Review, 11:1090–1098. **5, Grainger** et al. (2012), Science, 336:245–248. **6, Cornoldi** et al. (1981), Firenze: Organizzazioni Speciali. **7, Raven** (1949), London: Lewis & Co. and Harrap & Co. **8, Inhoff** & Rayner (1986), Perception & Psychophysics, 40(6):431–439. **9, Rayner** & McConkie (1976), Vision Research, 16:829–837.

European Research Council
Established by the European Commission