Early Interaction Effects of Word Type and Frequency on Lexical Decision Processes

Joshua Krause, Jelmer P Borst, & Jacolien van Rij
Bernoulli Institute, University of Groningen

university of groningen

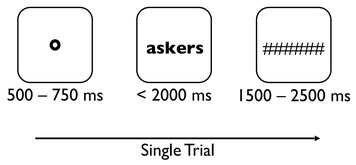
Introduction

- Previous work revealed that lexical decisions require completion of six cognitive processing stages[1]
- Lexical decisions for frequent and infrequent words differ only in duration of a **late** (~ 400 ms) processing stage[1]
- Using novel machine learning (HsMM-MVPA[2]) we investigated continuous frequency effects for words and non-words on **early** processing stages

Design & Methods

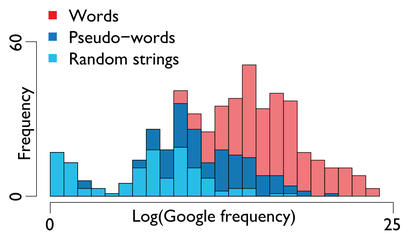
Lexical Decision Task

- 26 native Dutch speakers performed 500 lexical decision (LD) trials
- 125 pseudo-words, 125 random strings, 250 words
- Stimuli were obtained from the DLP[3] corpus



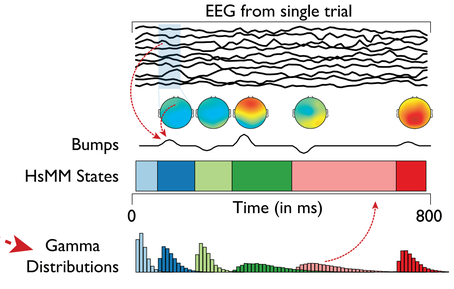
Google Frequency Measure

- For words and non-words the **Google result count**[4] was used as frequency measure
- Correlates with traditional frequencies but also other features[4] (e.g. OLD20)
- Reflects mixture of information - good **word familiarity measure**



Machine Learning (HsMM-MVPA)

- Recovers cognitive stages by finding stage-specific **bump** topologies in EEG (MVPA)
- Stage duration can differ between trials (HsMM)
- Provides **trial-level** cognitive stage onset and duration estimates[2]

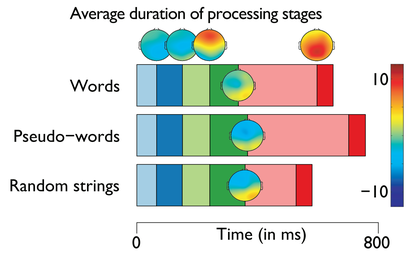


References:
[1] Berberyan et al., "Discovering the brain stages of lexical decision: Behavioral effects originate from a single neural decision process", Brain and Cognition, 2021
[2] Anderson et al., "The discovery of processing stages: Extension of Sternberg's method", Psychological Review, 2016
[3] Keuleers et al., "Practice effects in large-scale visual word recognition studies : a lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords", Frontiers in Psychology, 2010
[4] Hendrix & Sun, "A word or two about nonwords: Frequency, semantic neighborhood density, and orthography-to-semantics consistency effects for nonwords in the lexical decision task", Journal of Experimental Psychology: Learning, Memory, and Cognition, 2020
[5] Wood, Generalized Additive Models: An Introduction with R, Second Edition (2nd ed.)., 2017
[6] Keuleers et al., "SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles", Behavior Research Methods, 2010
[7] Hauk et al., "(Q) What Would You Prefer a SOSSAGE to a SAUSAGE? (A-) At about 100 msec. ERP Correlates of Orthographic Typicality and Lexicality in Written Word Recognition," Journal of Cognitive Neuroscience, 2006
[8] Segalowitz & Zheng, "An ERP study of category priming: Evidence of early lexical semantic access", Biological Psychology, 2009

Results

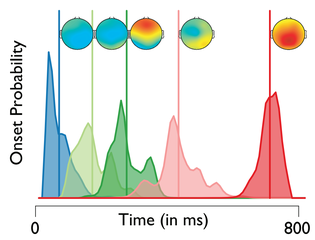
Word Type Has Main Effect on Late Processes (~ 300 ms)

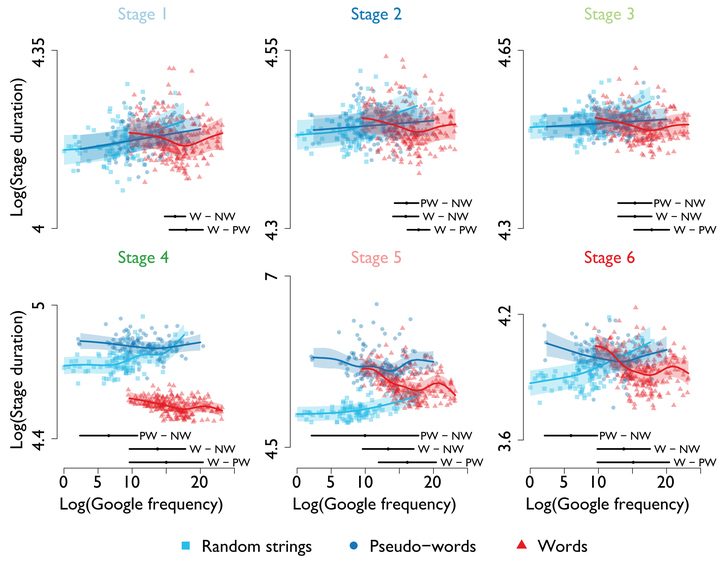
- Differences in duration of **fourth and fifth stage** between **word types**
- Scalp topology preceding fifth stage differs between word types



Frequency and Word Type Have Effects on Earliest Processes

- **Generalized additive mixed models**[5] (GAMM) were used to model word type and frequency effects
- **Weighted mean** of trial-level onset distribution estimated by HsMM-MVPA used as dependent variable

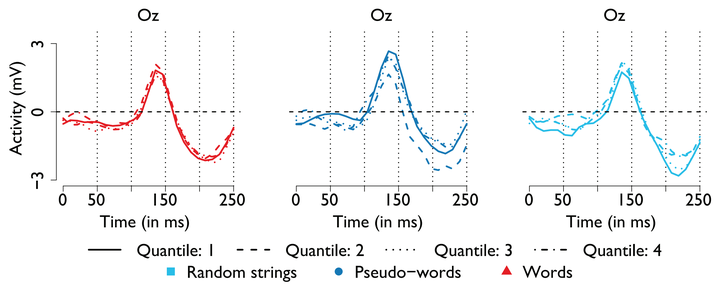




- **Surprisingly early** (< 100 ms) frequency and word type interaction
- Positive effect of frequency in early but negative effect in late stages
- **Prominent Word type effect** in stage four
- Complex **non-linear interaction effects** in stages five and six

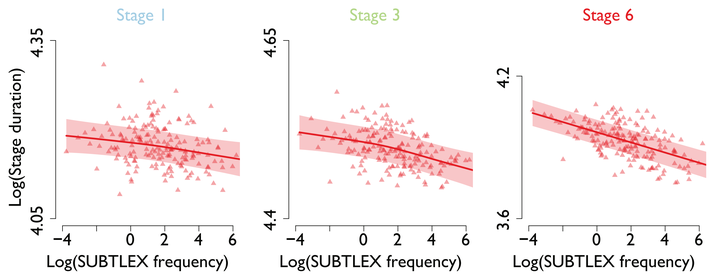
Discussion

Early Frequency Effects Robust to Alternative Analysis



SUBTLEX Frequencies Have Negative Effect on All Durations

- Different from Google frequencies where early positive effect reverses later
- SUBTLEX[6] does not reflect real-world **familiarity** effect on early processes



Limitation: GAMM analysis does not account for **onset uncertainty**

Conclusion

- Frequency effects in the first stages suggest **early visual processing**[7] and **orthographic encoding**[8] to be sensitive to stimulus properties
- Word type effect is very prominent in stage four - potentially indicating an early **binary discrimination** between words and non-words
- Non-linear interaction in stage five might suggest that more nuanced differences between stimuli are considered to **reach a conclusive LD**[1]

Word familiarity information is utilized differently at every processing step involved in LD depending on the role fulfilled by each step.

Reach out:
j.krause@rug.nl

Get the poster and abstract from my website!