Orthographic neighborhood density effects in a Maltese visual lexical decision megastudy

Introduction: The processing of real and non-word targets, as measured in visual lexical decision, is differently affected by changes in the target's orthographic neighborhood density (i.e. the number of words that differ from the target by the substitution, insertion, or deletion of a letter; e.g. *dot*, *dogs*, and *do* are neighbors of *dog*): Response times (RTs) decrease for real-word targets as neighborhood density increases, but increase for non-word targets owing to the activation of potential real-word candidates (Andrews 1989, Hendrix and Sun 2021). We analyze neighborhood density effects on the processing of real and non-words in Maltese, a Semitic language, using data from the MaltLex database of Maltese visual lexical decision responses (Geary 2020).

Maltese is unique among Semitic languages in that it is written using the Latin alphabet, while it is also unique among the languages studied previously in that it uses nonconcatenative morphology typical of Semitic languages (cf. Frost et al. 2005 likewise found a facilitatory effect of orthographic neighborhood density on responses to real-word targets in a Hebrew masked priming visual lexical decision study). Additionally, little research has explored the processing of Maltese non-words (cf. Twist 2006). Thus, our study explores orthographic neighborhood density effects in a language with a unique combination of orthographic and morphological properties, while also illuminating the processing of Maltese-like non-words by Maltese speakers.

MaltLex stimuli: 104 native Maltese speakers provided 237,000 lexical decisions to 11,040 real-word and 10,954 non-word targets in a series of Maltese visual lexical decision tasks (Geary 2020). Real-word targets were randomly selected from the Korpus Malti v3.0 corpus (Gatt and Čéplö 2013), and then checked against the online lexical database Ġabra (Camilleri 2013) and vetted by a native speaker. Non-word targets were constructed by replacing the consonant letters of a real-word target to make a phonotactically-legal non-word. Real- and non-word targets were matched in length (M = 7.1 letters) and frequency-weighted neighborhood density (M_{Real} = 157.4, M_{Nonce} = 123.9 occurrences per million), and a native speaker vetted all non-word targets.

Analysis: We analyzed 210,960 datapoints from visual lexical decision trials on which the participant provided the intended response to real-word and non-word targets (N_{Real} = 104,644, N_{Nonce} = 106,316 datapoints). We analyzed log RTs using the lme4 package (Bates et al. 2015) in R (R Core Team 2021) to fit an LMER model, with target lexicality (Real vs. Nonce; reference: Real), target log frequency-weighted neighborhood density, and the interaction of target lexicality and target log frequency-weighted neighborhood density as fixed effects. The model also included control variables like target length and participant's age as fixed effects, as well as subjects and targets as random effects. We used the lmerTest package (Kuznetsova et al. 2016) to simulate Satterthwaite approximations for degrees of freedom to assess the significance of fixed effects.

Results and Discussion: Consistent with previous research, the effect of target lexicality was significant ($\hat{\beta}$ = 0.014; $t(21{,}140) = 2.76$, $p < 0.01$), with participants responding slower to non-word targets (M = 990 ms) than real-word targets (M = 850 ms). The effect of neighborhood density was significant ($\hat{\beta}$ = −0.016; $t(21{,}220) = -25.78$, $p < 0.001$), with participants responding faster to real-word targets as neighborhood density increases, but the interaction of target lexicality and neighborhood density was also significant ($\hat{\beta}$ = 0.020; $t(20{,}800) = 28.482$, $p < 0.001$), with the facilitatory effect of neighborhood density diminishing for non-word targets. To investigate this interaction further, we split the dataset by target lexicality and re-fitted a model to each dataset: Consistent with previous research, the neighborhood density effect was facilitatory for real-word targets ($\hat{\beta}$ = −0.021; $t(10{,}430) = -30.914$, $p < 0.001$) but inhibitory for non-word targets ($\hat{\beta}$ = 0.010; $t(10{,}170) = 21.187$, $p < 0.001$), indicating that Maltese's unique orthographic and morphological characteristics do not alter the effects of orthographic neighborhood density on lexical processing.

References

Andrews, S. (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15: 802−814. https://doi.apa.org/doi/10.1037/0278-7393.15.5.802.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1−48. http://dx.doi.org/10.18637/jss.v067.i01.

Camilleri, J. J. (2013). *A computational grammar and lexicon for Maltese* (MSc Thesis). Chalmers University of Technology, Sweden. https://mlrs.research.um.edu.mt/resources/gabra/.

Frost, R., Kugler, T., Deutsch, A., and Forster, K. I. (2005). Orthographic structure versus morphological structure: Principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31: 1293−1326. https://doi.apa.org/doi/10.1037/0278-7393.31.6.1293.

Gatt, A., and Čéplö, S. (2013). Digital corpora and other electronic resources for Maltese. In *Proceedings of the international conference on corpus linguistics*. University of Lancaster.

Geary, J. (2020). MaltLex: A database of visual lexical decision responses to 11,000 Maltese words. Poster presented at *the 33rd Annual CUNY Conference on Human Sentence Processing* (CUNY 2020), Amherst, MA, March 20, 2020. https://osf.io/8rp3a/.

Hendrix, P., and Sun, C. C. (2021). 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* 47: 157−183. http://dx.doi.org/10.1037/xlm0000819.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2016). lmerTest: Tests in linear mixed effects models. [R package v. 2.0-32]. https://CRAN.R-project.org/package=lmerTest.

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Twist, A. E. (2006). *A psycholinguistic investigation of the verbal morphology of Maltese* (PhD Thesis). University of Arizona, Tucson, AZ. http://hdl.handle.net/10150/194996.