English Pseudoword Lexicon Project: a megastudy for investigating pseudoword processing

Fabio Marson[1], Rolando Bonandrini[1], Iva Šaban[1], Simona Amenta[1], Marco Marelli[1]

[1] *Department of Psychology, University of Milano-Bicocca, Italy*

Background: Research in psycholinguistics has been characterized by studies focused on properties of lexical items, with non-lexical items classically used as foils. Non-lexical items (i.e., pseudowords and nonwords) have been in fact neglected as a matter of research due the fact that they are not represented in the mental lexicon and because they have less properties of interests compared to lexical items (e.g., frequency or a specific referent). However, some studies show that orthographic features such as length and number of orthographic neighbors can affect their processing [1, 2], as well as the presence of a morphological structure [3, 4]. The current literature lacks a systematic exploration of how such features affect processing of non-lexical stimuli and, eventually, their effect on the lexical stimuli among which they are presented. The current project is aimed at filling this gap by providing a new megastudy designed to systematically investigates such effects and build a resource that researchers can consult to get prior data on non-lexical stimuli.

Method: Three types of non-lexical stimuli were adopted: Simple Pseudowords (SPW; unattested strings of letters that follow English graphotactics generated by Wuggy, e.g., "*plief*", N = 4523), Complex Pseudowords (CPW; unattested strings of letters that follow English graphotactics generated by concatenating existing stems and suffixes, e.g., "*plungable*", N = 4498), and Nonwords (NW; strings of letters that do not follow English graphotactics generated by concatenation of letters, e.g., "*iprwvca*", N = 4475). These stimuli were presented along with existing Words, half of which were morphologically simple (not suffixed) and half complex (stem + suffix). We adopted a lexical decision task to obtain RTs for all stimuli. Crucially, we divided the data collection in four different studies across which the lexical stimuli remained the same, but the type of non-lexical stimuli differed by design: SPW, CPW, and NW were administered in three separate studies (Exp1, 2, 3), with a final study in which all the three types were used (N = 1500 for each type; Exp4). The data collection was conducted online involving 1416 English (US) native speakers.

Results: Non-lexical items were found to be overall processed faster than words (mean RTs = 749 vs 770 ms, respectively). More in detail, in experiments designed to have only one type of non-lexical items, CPW were found to be the slowest (mean RTs = 860 ms), followed by SPW (774 ms), and then NW (579 ms). The same pattern with overall slower RTs was found when all types of non-lexical items were used in the same experiment (CPW = 954 ms; SPW = 786 ms; NW = 602 ms), with CPW showing the greatest difference (ΔRTs = 94 ms). Also, RTs for words were slower whenever CPW were present in the list (Exp2 = 819 ms; Exp4 = 797 ms), compared to the exclusive presentation of SPW (Exp1 = 786 ms), and NW (Exp3 = 579 ms). Finally, we observed that RTs for simple and complex words were differently modulated by the non-lexical items they were presented with, showing that responses for complex words were slower whenever CPW were presented.

Discussion: Our results showed that RTs for non-lexical items are affected by morphological and orthographic features, as well as by list effects. Summing up, we observed a sort of progressively increased difficulty in rejecting non-lexical items with the increase of their sublexical organization, from the less structured (NW) to the more structured (CPW). Finally, we report that even the processing of words is affected by the non-lexical items that accompany them, thus advising more cautious stimuli selection when deciding the study design in psycholinguistic research. Our investigation provided a comprehensive resource that researchers can use to carry on additional analyses to test hypotheses or to improve stimuli selection for their studies and have more control over experimental designs involving non-lexical stimuli.

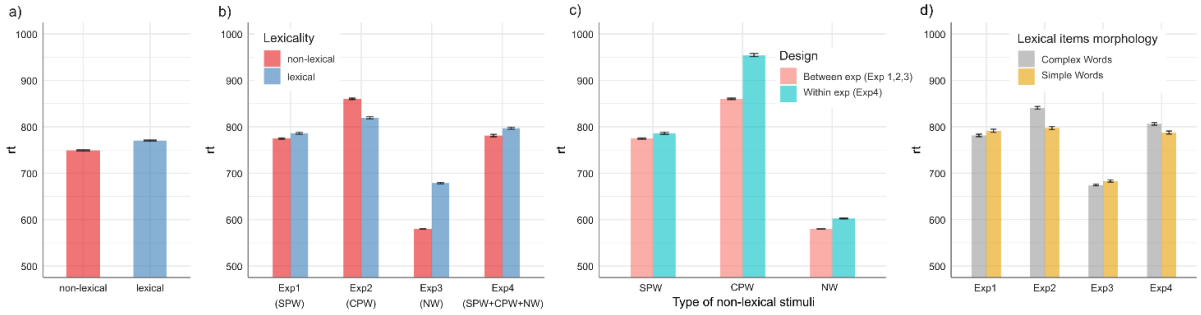

Figure 1: Barplots representing RTs and standard errors for lexical vs non-lexical items averaged across all experiments (a), and lexical vs non-lexical items separated for each experiment (b). Third panel (c) represents RTs and standard errors for the three types of non- lexical items when presented as the only type in the list (Between exp condition) or when presented with other non-lexical items (Within exp condition). The last panel (d) shows RTs and standard errors for simple and complex words in different experiments.

References

[1] Yap, M., Sibley, D., Balota, D., Ratcliff, R., & Rueckl, J. (2015). Responding to nonwords in the lexical decision task: Insights from the English Lexicon Project. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 597–613. https://doi.org/10.1037/xlm0000064

[2] Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., & Yap, M. (2004). Visual Word Recognition of Single-Syllable Words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

[3] Taft, M., & Forster, K. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, *14*(6), 638–647. https://doi.org/10.1016/S0022-5371(75)80051-X

[4] Crepaldi, D., Rastle, K., & Davis, C. (2010). Morphemes in their place: Evidence for position-specific identification of suffixes. *Memory & Cognition*, *38*(3), 312–321. https://doi.org/10.3758/mc.38.3.312