Covert structure and zero morphology in sentence processing

Alexandra Krauska & Masaya Yoshida (Northwestern University)
AlexKrauska2019@u.northwestern.edu

A word such as "visit" can be used as a verb (*I expect to **visit***) or as a noun (*I expect a **visit***) due to the presence of unpronounced morphology, or "zero morphology."[1,2,3,4] Past studies have uncovered some grammatical properties of these categorically ambiguous words, but no research has been conducted on how people process these words in reading. Are these two different uses of "visit" processed in the same way? Does covert structure have an effect in online sentence processing? Through an eye-tracking experiment, this study shows that the zero-derived form of categorically ambiguous words results in a slower reading time, which we argue is due to the structural complexity of those words.

It has been suggested that the different forms of categorically ambiguous words are morphologically derived from their base form. Under such analysis, *visit* as a noun ([$_N$ visit]) is derived from visit as a verb ([$_V$ visit]). This derivation process involves an invisible zero-suffix as in [[visit $_V$] ø $_N$] [1,2,3,4]. It has been long known that words and sentences that have more complex structure incur more processing difficulty than those which have simpler structure[5,6]. Therefore, if zero-derived words have a more complex structure than non-zero-derived words, then we expect the processing of zero-derived words to induce greater costs than their root counterparts. Furthermore, it has been argued that the mechanism of sentence processing relies on information that we can directly observe, such as words or explicit suffixes[7,8]. If zero-derived words are indeed processed slower than non-zero-derived words, it means that the mechanism of sentence processing can recognize the invisible suffix, and that structures of sentences and words can be built based on the invisible element.

In this study, an experiment using eye-tracking while reading was conducted (n=48) in which Derivation x Type of Derivation (N → V: *paint* vs. V → N: *visit*) were manipulated as independent factors in a 2x2 factorial design as in (1). We developed the 24 stimuli using categorically ambiguous words investigated in [9], those which the verb/noun COCA[10]-based frequency in both the derived and underived conditions, acceptability, semantic similarity, and argument structure were tightly controlled ($p > .05$). The base category of each word was determined using category-specific root morpheme acceptability (such as: *Darwin-ian-ism/ *Darwin-ism-ian*). Each sentence was rated for acceptability beforehand (n = 80).

(1) a. John expected the paint behind his workshop. *Non-derived noun*
    b. John expected to paint behind his workshop. *Derived verb*
    c. John expected to visit after the doctor called. *Non-derived verb*
    d. John expected the visit after the doctor called. *Derived noun*

Linear Mixed Effects Regression models of the data with derivation as a predictor, and item, subject, and acceptability judgement rating as fixed effects (*see graph*) revealed a significant main effect of Derivation on log reading time at the critical region (paint/visit). In the Derived conditions (1b/d), *paint/visit* was read significantly slower than in the Underived conditions (1a/c) in first pass reading time, regression path reading time, and total reading time at the critical region ($p < 0.05$). Other measures and other regions did not reach significance. There was some interaction between Derivation x Base Category. This experiment suggests that the slower reading time for the derived words is a result of the structural complexity of zero-derivation. This result is not predicted from an analysis of categorically ambiguous words where no morphological derivation is involved or in which morphological complexity does not impede sentence comprehension. Furthermore, this result supports the claim that the sentence processing mechanism is sensitive to the presence of invisible suffixes, and that structure is built even without visible morphological elements.

First Pass and Regression Path
Log Reading Times at the Critical Region
Controlling for Subject, Item, and Acceptability Rating
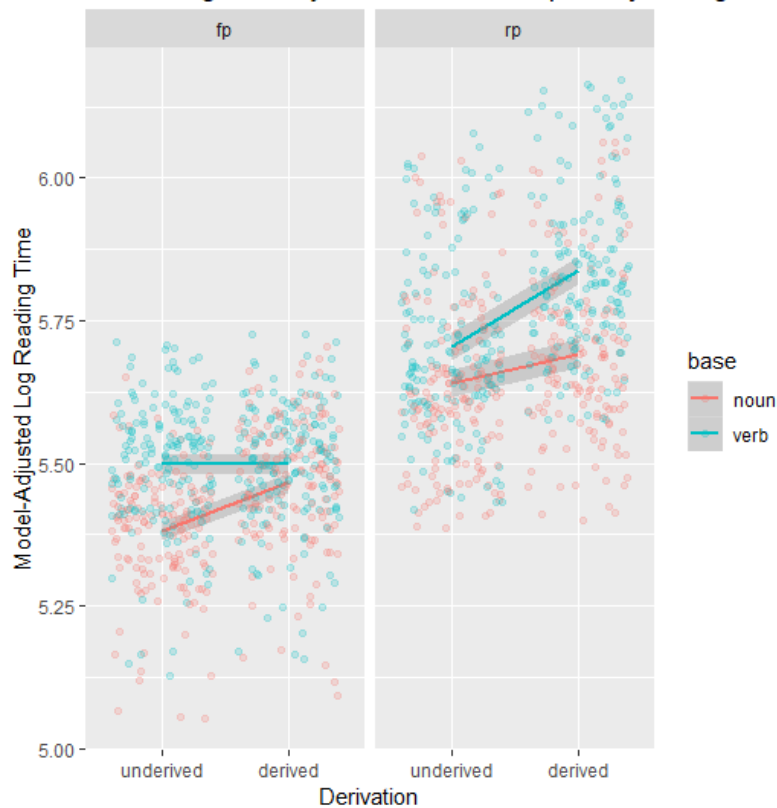
References:

[1] Clark, H. H., & Clark, E. V. (1979). When nouns surface as verbs.
[2] Marchand, H. (1964). A set of criteria for the establishing of derivational relationship between words unmarked by derivational morphemes.
[3] Myers, S. (1984). Zero-Derivation and Inflection. *Papers from the January 1984 MIT Workshop in Morphology, MIT Working Papers in Linguistics*.
[4] Pesetsky, D. (1996). *Zero syntax: Experiencers and cascades*. Cambridge: Mit Press.
[5] Miller, G. A., & Chomsky, N. (1963). Finitary models of language users.
[6] Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1-76.
[7] Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. MIT Press.
[8] Barry, G., & Pickering, M. (1990). Dependency and constituency in categorial grammar. *Barry, G. and Morrill, G*.
[9] Lukic, S. (2016) Neurocognitive Correlates of Nouns and Verbs: Zero-Derivation and Lexico-Semantic Processes (Doctoral Dissertation, Northwestern University)
[10] The corpus of contemporary American English. http://corpus.byu.edu/coca/.