Getting to the 'root' of semantic and syntactic processing of morphologically complex words: MEG evidence from Arabic

Suhail Matar[1,2], Samantha Wray[2,3], Sherine Bou Dargham[4], Linnaea Stockall[2,5], and Alec Marantz[2,4]

[1] The Basque Center for Cognition, Brain, and Language, Donostia–San Sebastián, Spain
[2] New York University, New York, NY, USA
[3] Dartmouth University, Hanover, NH, USA
[4] New York University Abu Dhabi, Abu Dhabi, UAE
[5] Queen Mary University of London, London, UK

Crosslinguistic research shows that syntactic and semantic processing during word reading recruits different neural bases at distinct latencies. This dissociation is exemplified in reading morphologically complex pseudo-words with one of two types of ill-formedness: a mismatch between an affix and a stem's syntactic category (e.g. *re-soft, cf. resend), or a semantic incompatibility between stem and affix (e.g. *re-smile). Semantic incompatibility has been found to elicit greater activity in orbitofrontal cortex compared to syntactic mismatch 350-500ms after word onset (see Stockall et al. (2019) for English and Neophytou et al. (2018) for Greek).

In languages which primarily use concatenative morphology (e.g., English), affixes are appended to cohesive, contiguous stems. But in Arabic and other Semitic languages, a major morphological hallmark is the interleaving of noncontiguous roots (e.g. {k,-t,-b}, indicating the lexical domain of 'writing') with noncontiguous templates (e.g., ma***CC***a***C***, a location template, where the *C*s indicate ordered root consonants) to build words ('ma***kt***a***b***' = 'office'). Arabic roots are traditionally analyzed as being underspecified; a root may appear in verbs more often than nouns, but does not receive either its category or full semantic information until after interleaving with a pattern. Thus, using the above ill-formedness paradigm in Arabic word reading allows us to determine if the separability of semantic and syntactic processing found in previous literature is actually due to more underspecified distributional properties.

N=18 Arabic speakers participated in a visual lexical decision task with concurrent magnetoencephalography. In addition to reading grammatical, attested words, speakers read two types of nonwords: Syntactic Violation, comprised of roots attested only in nominal templates, interleaved with a verbal template (e.g. {ʕ,q,r,b} 'scorpion' + ta***C***a***CC***a***C*** 'passive/reflexive' = '*scorpioned'); and Semantic Violation, comprised of roots attested in verbs but interleaved with a passive/reflexive verbal template in which they are unattested (e.g. {z,ɣ,r,d} 'trill' + ta***C***a***CC***a***C*** 'passive/reflexive' = '*was trilled'). Spatiotemporal cluster-based regressions at 10,000 permutations (following Maris & Oostenveld 2007) were performed to determine clusters of differences in brain activity in time and space within a window of 300-500ms after stimulus presentation. We identified two clusters in orbitofrontal cortex: in the left hemisphere, Semantic Violation items elicited significantly more activity than Syntactic Violation items from 426-492ms ($p = 0.052$; Fig. 1); in the right hemisphere, Semantic Violation items elicited more activity than Syntactic Violation items (421-478ms, $p = 0.017$; Fig. 2).

These results suggest that the processing stages of morphologically complex words which have previously been claimed to be sensitive to syntactic category (Stockall & Manouilidou 2014;

Schreuder & Baayen 1995) may actually be picking up on more distributed statistical properties of morphosyntax, such as which syntactic category is the root/stem likely to emerge in. These results thus not only inform us about how syntactic and semantic information is processed during word recognition but is also compatible with theoretical linguistic accounts that posit distribution of semantic and syntactic properties across the grammar such as Distributed Morphology (Halle & Marantz 1994), as opposed to the bundling of these properties within lexical items.

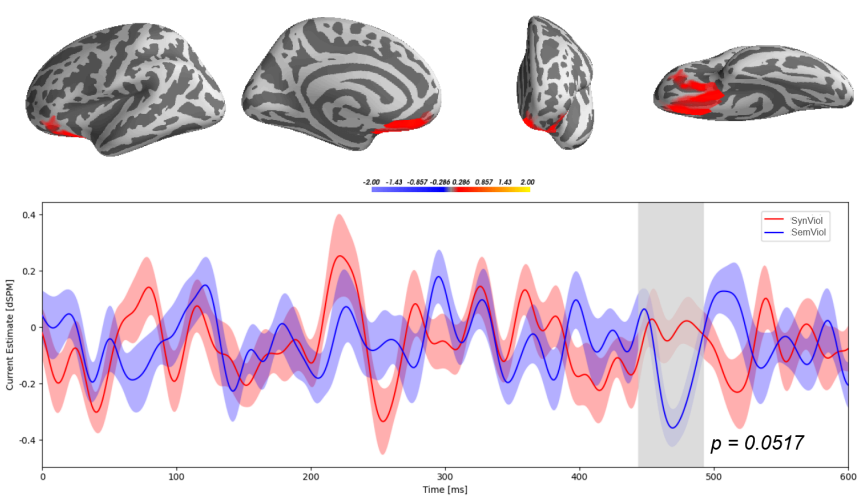

Figure 1: *Top: location of cluster in left hemisphere showing increased activity (higher negative value) for words with Semantic Violations of compatibility between their root and verbal template. Bottom: time course of activity in indicated area for both Semantic Violation and Syntactic Violation items, with temporal cluster indicated in gray.*

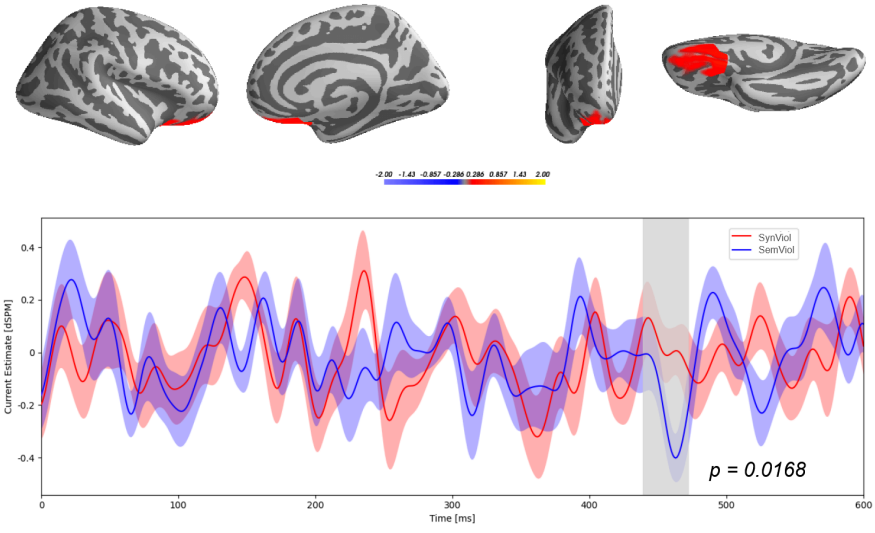

Figure 2: *Top: location of cluster in right hemisphere showing increased activity (higher negative value) for words with Semantic Violations of compatibility between their root and verbal template. Bottom: time course of activity in indicated area for both Semantic Violation and Syntactic Violation items, with temporal cluster indicated in gray.*

References

Halle, M., & Marantz, A. (1994). Some key features of Distributed Morphology. *MIT working papers in linguistics*, *21*(275), 88.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of neuroscience methods*, *164*(1), 177-190.

Neophytou, K., et al. (2018). Syntactic and semantic restrictions on morphological recomposition: MEG evidence from Greek. *Brain & lang*, *183*, 11-20.
Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. *Morphological aspects of language processing*, *2*, 257-294.
Stockall, L., & Manouilidou, C. (2014). Teasing apart syntactic category vs. argument structure information in deverbal word formation: a comparative psycholinguistic study. *Italian J of Ling*.
Stockall, L. et al. (2019). Prefix stripping re-re-revisited: MEG investigations of morphological decomposition and recomposition. *Frontiers in Psychology*, *10*, 1964.