Studying the effects of non-concatenative morpho-semantic & morpho-syntactic anomalies in Modern Standard Arabic

Yasmin Beydoun, Kaitlyn Tagarelli, & Aaron J. Newman

Modern Standard Arabic (MSA) is a literary standard used across the Middle East and is generally reserved for writing and formal speaking occasions. MSA is a derivational and nonconcatenative language, meaning a majority of its words are composed of overlapping morphemes derived from a common 3-consonant root. In English, its morphemes are strung together sequentially, such as for example *walk-ed*. MSA is composed of two morphemes, semantic and syntactic, which are overlapped to form new words. The semantic morpheme consists of three consonant roots that carry the meaning of the word. For example, *KTB* (pronounced *kataba*) is a root meaning "he wrote". The syntactic morphemes, or word patterns, are inserted into the roots to produce new words. Examples of this are the word patters *–i-a-* and *-aa-i-*, which are inserted to create *kitab* and *kaatib* (*"*book" and "male writer," respectively). What is interesting to note is that these words are semantically related to the original root, meaning they belong to a single lexical field. While less common, it is possible to create words that are not semantically related to their root, such as for example *iktitaab*, meaning "registration".

Although MSA possesses a rich grammar that differs significantly from other languages, it has been understudied using neuroimaging. The proposed project aimed to study sentence processing in MSA. The objective of the study was to better understand how a language with overlapping morphology is processed — specifically, whether neurocognitive distinctions between semantic and syntactic word processing are the same in languages with overlapping morphology as with sequential morphology. As MSA has different morphemes overlapping to create words, it is of interest to understand the different types of processing associated with each morpheme. To fulfill the objective, the following question was asked: Will anomalies of semantic and syntactic morphemes in MSA elicit different EEG responses, specifically the N400 and P600, respectively?

To explore this question, we tested native Arabic speakers. The conditions were in the form of identical sentences with the exception of the anomaly found at the end. There were four conditions, in which a target word was either: correct, a semantically unrelated morpho-semantic violation, a semantically unrelated morpho-syntactic violation, or a semantically related morpho-syntactic violation. As participants were presented with sentences on a computer screen, one word at a time, they were also asked to judge (behavioural measure) how good or bad the sentence was via a Likert scale (1 to 5). The task was performed while EEG was recording.

The EEG data were collected using a 16-channel ActiCAP system, and focused on the analysis of two ERPs: the N400 associated with semantic violations and the P600 associated with syntactic anomalies. The EEG and behavioural data were used as the main measures of the study. Results have implications for better understanding sentence processing in under-studied languages.