ELSEVIER

# Morphological representations are extrapolated from morpho-syntactic rules

L. Gwilliams[a,c,*], A. Marantz[a,b,c]

[a] Department of Psychology, New York University, United States
[b] Department of Linguistics, New York University, United States
[c] NYUAD Institute, New York University, Abu Dhabi, United Arab Emirates

ABSTRACT

The field of psycho- and neuro-linguistics has long-debated the decompositional model of visual word processing: Are written words processed via the visual forms of stem and affix morphemes, or as complex wholes? Although many have now settled upon a decompositional view, it is unclear what heuristic the brain uses to generate these visual morpheme-forms in the first place. Here we conduct a magneto-encephalography study to test two hypotheses for how this may be done: i) the brain encodes representations of the morphemes that follow the morpho-syntactic rules governing constituents: A stem morpheme will be represented if the word obeys the grammatical behaviour associated with its suffix; ii) the brain only encodes stem morphemes that occur with multiple suffixes or as words in isolation. Our results indicate that words with morpho-syntactic wellformedness as stem-suffix combinations are decomposed by the system, thus supporting the former hypothesis. This suggests that knowledge of morpho-syntactic rules can be used to form morphological representations of written words, in absence of independent experience with all of their constituent morphemes. Possible mechanisms supporting this computation are discussed.

## 1. Introduction

Central to understanding human communication is identifying the building blocks of language. Which linguistic units are committed to memory and subsequently retrieved? How are words, which vary along a number of orthographic, phonological, syntactic and semantic dimensions, represented in the brain?

The role of morphological structure for the organisation of the mental lexicon has been a heated topic for almost half a century, debating whether words are represented in terms of constituent morphemes (e.g., {farm}, {-er}) or whole words (e.g., {farmer}). A large body of research has demonstrated that the visual system indeed processes words through constituent morphological representations. This has been shown using a variety of behavioural methodologies, and across a number of different languages (Taft and Forster, 1975; Taft, 1979; Caramazza et al., 1988; Colé et al., 1989; Grainger et al., 1991; Marslen-Wilson, 1994; Deutsch et al., 1998; Rastle, 2000; Clahsen, 2003; Rastle et al., 2004; Longtin and Meunier, 2005; Duñabeitia et al., 2007; Meunier and Longtin, 2007; Rastle and Davis, 2008; Diependaele et al., 2009; Gwilliams et al., 2015), though see (Pinker and Prince, 1988; Giraudo and Grainger, 2000; Seidenberg and Gonnerman, 2000; Pastizzo and Feldman, 2002; Feldman, 2004; Gonnerman et al., 2007; Baayen, 2011) for a non-decompositional stance on visual word processing.

Sensitivity to morphological structure has also been corroborated in neurophysiological work. Responses in the fusiform gyrus are modulated by the visual complexity of a word; one important aspect of this complexity being how many morphemes it contains (Pylkkänen, 2004; Lavric et al., 2007; Morris et al., 2008; Zweig and Pylkkänen, 2009; Solomyak and Marantz, 2010; Lehtonen et al., 2011; Lewis et al., 2011; Morris et al., 2013; Fruchter and Marantz, 2015; Whiting et al., 2015; Cavalli, 2016). For example, in a magneto-encephalography (MEG) study, Zweig and Pylkkänen (2009) identified a response component that elicited increased activity for bi-morphemic words (e.g. *farm-er*) as compared to both monomorphemic orthographic controls (e.g. *wint-er*) and simple monomorphemic words (e.g. *switch*). This was observed ~170 ms after visual word presentation in the fusiform gyrus; corresponding to the timing and location of the M170 response (Pylkkänen and Marantz, 2003). Conditions were matched along a number of dimensions known to affect lexical processing, such as length, surface frequency, lemma frequency and orthographic neighbourhood frequency. This result suggests that there is a measurable brain response to visual complexity in the fusiform gyrus — a neural indicator of how many "parts" make up a visual object.

The location of this M170 response in the left fusiform can be linked to the putative visual word form area (VWFA). This corresponds to a region in the left lateral occipito-temporal sulcus that is activated in response to written words, regardless of location (Cohen, 2000) and

---

* Corresponding author at: 10 Washington Place, 6th Floor, New York, NY 10003, USA.
E-mail address: laura.gwilliams@nyu.edu (L. Gwilliams).

case (Dehaene and Naccache, 2001). This region is thought to be sensitive to the abstract sequence of letters that compose a written string — the visual word form. Finding that this region also responds to morphological complexity suggests that it may be recognising the "visual morpheme-forms" contained within the written word.

Furthermore, previous research has demonstrated that the M170 response is modulated by the transition probability (TP) between the morphemes of a bi-morphemic word (Solomyak and Marantz, 2010; Lewis et al., 2011). TP quantifies the probability of a suffix (e.g. -er) attaching to a certain stem morpheme (e.g. farm) given all the possible suffixes that could attach to the stem (e.g. -s, -er, -ing, -ed). This measure specifically refers to the transition between morphological units, not simply phoneme/letter strings. Finding a correlation between neural responses and this variable suggests that the brain tracks the statistical regularity between morphological constituents. Whether or not the brain is sensitive to the transition probability between morphemes may therefore be used as a index of whether decomposition has occurred.

So, it has been established that visually presented words are processed via constituents, and there is a brain response located in the fusiform that is sensitive to the morphological complexity of a written word. Building on these results, a critical question arises: How does the processing system recognise that a word is indeed morphologically complex? One way to address this is to test which words the system decomposes, and which it does not.

Previous studies have shown that not just free stem words are decomposed (e.g. farmer → farm), but also bound stem (e.g. sociable → social), and irregular forms (e.g. fell → fall) (Crepaldi et al., 2010; Fruchter et al., 2013). Further, pseudo-complex words like *brother* and *corner* appear to be decomposed into *broth + er; corn + er*, at least in the initial stages of processing, but visually similar words without a suffix are not (e.g. broth-el) (Rastle et al., 2004). Words containing a pseudo-suffix (e.g. wint-er) are also not decomposed (Zweig and Pylkkänen, 2009). What do these results suggest? i) The morphological parser seems to be robust against the kind of orthographic alternations found in bound stem items, as well as infrequent irregular derivations. Likely, then, the parser is dealing with abstract representations that surpass simple visual template matching. ii) It is unclear whether the semantic relationship between the stem and whole word factors into the initial decompositional process (cf. see Diependaele et al., 2009; Devlin, 2004; Feldman et al., 2009) for the debate on blind decomposition within the masked priming literature, though we will not delve further into this here). iii) The presence of a stem or suffix morpheme is necessary but not sufficient to initiate a decompositional parse (because neither *broth-el* nor *wint-er* are decomposed). This suggests that the system does not just decompose based on the recognition of a stem or suffix alone, but also requires that the word meets an additional criterion. Here, we want to identify that criterion.

## 1.1. Aims

Based on these previous results, we come to the current question: *how* are visual morpheme-form representations created in the first place; what heuristic does the visual system use to recognise a morphological constituent as such, and commit it to memory?

In order to explain the different predictions of the hypotheses we will test, it is important that the reader understands the difference between two critical types of items. First are "excursion"-type words. These items contain a string sequence that matches a suffix in English (e.g. -ion), and critically they <u>do</u> behave in line with that suffix's function — *excursion* could be the de-verbal noun derived from the verb *excurse*. We will refer to these words as *valid-rule* because, based on the morpho-syntactic rule ascribed by the suffix, one can recover the stem "excurse" after being exposed to the complex word "excursion". Second are "winter"-type words, as tested by Zweig and Pylkkänen (2009). These items also contain a string sequence that matches a suffix (e.g.

-er), but they do not behave in line with the function typically ascribed by that suffix — winter is not an agentive noun that could be paraphrased "something that wints". We will refer to these types of words as *no-rule* because the potential stem "wint" cannot be recovered from the putative complex form "winter". The probability of transitioning from potential-stem to potential-suffix is equal to 1 for both word types; so, the critical difference between these conditions is whether or not the word behaves in line with the morpho-syntactic rules of the suffix.

With this in mind, we test the following two hypotheses.

**Hypothesis 1.** Visual morpheme-form representations are only generated when a stem morpheme is encountered in particular lexical contexts: either in isolation as words (e.g. farm) or within a number of morphologically complex words (e.g. farmer, farming, farmed). This hypothesis predicts that neither the no-rule *winter* nor the valid-rule *excursion* words are decomposed, because neither *wint-* nor *excurse-* occur in any other context than with the potential-suffix *-er* and *-ion*, respectively. Computationally speaking, then, any (and only) a word for which the probability of transitioning from stem to suffix is less than 1 will be represented in terms of its constituent morphemes. If this is the case, we would expect to find that activity in the fusiform gyrus ~170 ms responds indistinguishably to *no-rule* and *valid-rule* words, and that both are significantly different from words that contain isolatable stem morphemes.

**Hypothesis 2.** Representations are formed based on the morpho-syntactic *rules* that govern constituent morphemes. From this perspective, a representation of the stem *excurse* would be generated because it obeys the morpho-syntactic rule enforced by the de-verbalising suffix *-ion* in the word *excursion*. However, no-rule words like *winter* would still be represented and processed as unanalysed wholes because there is no morpho-syntactic rule supporting their decomposition. This is different from the first hypothesis in the critical aspect that the transition probability need not be less than 1 in order for a word to be compartmentally represented — the only thing that matters is valid morpho-syntactic structure. If this is the case, we would expect to find that *no-rule* and *valid-rule* words elicit significantly different M170 responses, and that *valid-rule* words are indistinguishable from words with an isolatable stem.

The aim of the present study is to adjudicate between these two hypotheses. Both predict that truly complex words (like *farmer*) are decomposed and no-rule *winter*-type words are not; however, they make different predictions about where valid-rule *excursion* words fall relative to these two word-types. The question then is simply whether brain responses to words with a valid morpho-syntactic structure pattern more closely with the truly complex words (suggesting that they are considered visually complex) or more closely with the simple words (suggesting that they are not considered visually complex).

In order to test this, we fit statistical models to explain neural responses in the left fusiform gyrus, during the time-window associated with morphological processing — the M170.

## 2. Method

A portion of the MEG data analysed here is reported as part of a different study, which primarily aimed to address methodological concerns regarding MEG source localisation (Gwilliams et al., 2016). Concretely, in the previous study, three datasets were used: one to replicate a previous finding (Experiment 1); one to generate a functional localiser of orthographic and lexical processes (Experiment 2); one to validate the accuracy of that localiser (Experiment 3 — a subset of the present data). We found that when using the data from Experiment 2, we could accurately identify a brain region (within the fusiform gyrus) that was sensitive to transition probability for the truly complex and pseudo complex items of the present study, replicating (Solomyak and Marantz, 2010; Lewis et al., 2011).