

## Morphological representations are extrapolated from morpho-syntactic rules

L. Gwilliams<sup>a, c, \*</sup>, A. Marantz<sup>a, b, c</sup>

<sup>a</sup> Department of Psychology, New York University, United States

<sup>b</sup> Department of Linguistics, New York University, United States

<sup>c</sup> NYUAD Institute, New York University, Abu Dhabi, United Arab Emirates

### ARTICLE INFO

**Keywords:**

Morphology

Visual word recognition

Neural representations

Magnetoencephalography

### 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.

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

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 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 occur with 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 an 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 do behave in line with that suffix’s function — *excursion* could be the de-verbal noun derived from the verb *excuse*. 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 “excuse” 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 *excuse-* 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 *excuse* 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).

### 2.1. Participants

Twenty four native English speakers from the NYUAD community took part in the study and were compensated for their time. 16 were female, aged between 19 and 50 ( $M = 21.9$ ,  $SD = 6.18$ ). All participants provided written informed consent prior to the experiment.

### 2.2. Materials

Four experimental conditions were created. Two conditions comprised of words that we expect to be decomposed, based on previous studies. Namely, truly complex words (e.g. builder, sociable) and pseudo-complex words (e.g. corner, brother). The *truly complex* condition contained 106 words: 53 free stem and 53 bound stem; the *pseudo-complex* condition contained 53 words. The other two conditions were the critical items in distinguishing the hypotheses outlined in the introduction: We gathered 53 *winter*-type words, for which there is no morpho-syntactic rule that could be used to generate the stem “wint” (*no-rule*), and 53 *excursion*-type words, for which there is a valid morpho-syntactic rule that could be extrapolated to derive the stem “excuse” (*valid-rule*).

We assigned words to conditions in the following way. First we extracted all items that were coded as mono-morphemic in both the English Lexicon Project (ELP) (Balota, 2007) and CELEX (Baayen et al., 1993), which also contained any of the suffix strings [*able*, *age*, *al*, *ant*, *ard*, *ate*, *er*, *et*, *ey*, *ic*, *ion*, *ity*, *let*, *ous*, *ic*, *ry*, *ward*]. This yielded 386 items. Next, we manually placed the words into one of three conditions: pseudo-complex (e.g. corner), valid-rule (e.g. excursion) and no-rule (e.g. winter). Pseudo-complex words contained a morphologically unrelated free-stem that attached to the suffix. The valid-rule items were compatible with the meaning and grammatical category of an extant morphologically complex word with the same suffix, while the no-

rule items differed in meaning and/or grammatical category from items derived with the same suffix. For example, “winter” is neither an instrument nor a person associated with a verbal meaning (a “winter” cannot be conceived as a thing used for “winting” or a person that “wints”), while “excursion” is plausibly the result of “excusing”. To determine which condition an item belonged to, we used exactly this kind of semantic/syntactic test, comparing the item at hand (e.g. “winter”) to a word that uses the same suffix in a typical morpho-syntactic context (e.g. “freezer” — something that freezes, “baker” — someone who bakes). If a suitable analogy was found, for example, between “explosion” — the result of exploding, and “excursion” — the result of excusing, it was deemed a valid-rule item; else, it was deemed a no-rule item.

We want to be the first to point out that there is a level of subjectivity in the test used to assign items to conditions, because the semantic and syntactic function of suffixes is not always clear. However, critically, there is no difference between the valid-rule and no-rule conditions other than what we can glean from the intuitive test we applied before running the MEG experiment. The conditions were matched along a number of linguistic dimensions that have been found to be relevant to lexical processing (see Table 1). Should the reader find an alternative characterisation that divides the no-rule and valid-rule classes, follow-up experiments should separate our account of the difference from the proposed alternative. But, to the best of our knowledge, the only manipulation between the critical conditions is whether or not a valid morphological rule exists. Therefore, we are forced to conclude that our results are driven by the intended manipulation, and not a confounding variable.

As a final point, although the factorial nature of this design forced us to make a binary distinction between “valid-rule” and “no-rule,” there is no argument from linguistic analysis that speakers are forced to make a binary decision, rather than, for example, weighting each option probabilistically. Therefore, our criteria for separation may be selecting items at extreme ends of what is really a continuum of items. Within the continuum are all words that end in orthographically English suffixes, thus qualifying as candidates for analysis in terms of a morphosyntactic rule. Further research on English derivational suffixes might allow us to evaluate “valid-rule” as a continuous variable.

Truly complex words were selected by extracting all words that were coded as containing two morphemes in both the ELP and CELEX. This yielded a pool of 371 words.

After identifying potential items for the four conditions, we then selected words that allowed the conditions to be matched on a range of stimulus properties. The distribution of suffixes over conditions was matched as closely as possible. Valid word items were matched across conditions for stem and word length, log mean bigram frequency, log orthographic affix frequency (frequency of the letter sequence making

**Table 1.**  
Summary of lexical statistics across experimental conditions.

Condition	Length stem (letters)	Length word (letters)	Log bigram frequency	Log orthographic affix frequency	Log morphological affix frequency	Log surface frequency	Log lemma frequency	Log transition probability
Free stem ( <i>Builder</i> )	4.91	7.19	3.63	5.51	4.68	1.8	3.08	1.28
Bound stem ( <i>Sociable</i> )	4.32	7.19	3.53	5.07	4.45	1.75	2.49	.75
Pseudo- stem ( <i>Corner</i> )	4.71	7.19	3.54	5.16	4.1	1.65	2.78	1.29
Valid rule ( <i>Excursion</i> )	4.56	7.19	3.59	5.16	4.13	1.69	NA	NA
No rule ( <i>Winter</i> )	4.72	7.19	2.54	5.31	4.22	1.61	NA	NA

up the affix), log morphological affix frequency (frequency with which the affix is used as a valid morpheme) and log surface frequency ( $p > .5$ ). Log lemma frequency and log transition probability from stem-suffix was also matched for between the free-stem conditions (summarised in Table 1). Keeping a balance between items-per-condition and how well matched the conditions were, resulted in 53 items per condition.

265 non-words were also created, in order to form the lexical decision task. All non-words contained letter sequences that match a valid English suffix. Sixty-five of the non-word items had grammatical well-formedness with a free stem (e.g., *drinkage*) and 65 were grammatically malformed with a free stem (e.g., *finalic*). The remaining items were formed from either a bound stem or a monomorphemic word (e.g., *medital*, *humilier*) and served as fillers.

### 2.3. Experimental procedure

Participants completed a visual lexical decision task on the 530 items. The order of items was randomised for each participant, and was split up into five blocks to avoid effects of fatigue. Responses were made with the left hand on a button-box while lying down in the MEG machine and looking up to the screen positioned around 85 cm away from the individual's face.

Each trial began with a fixation cross for 400 ms before the presentation of the stimulus. Participants had 2000 ms to indicate via button press whether the item was a valid word of English. A blank screen would then appear for 300 ms before the next trial began.

### 2.4. Data collection

Each participant's head was digitally scanned with a hand-held FastSCAN laser scanner (Polhemus, VT, USA), along with five digital points: three points on the forehead and one either side of the left and right auditory canal. Five electro-magnet marker coils were placed in the same position that had been digitally marked on the scan, and their position relative to the MEG sensors was recorded just before and just after the experiment. Together, the digital headshape, points and marker coil positions were used to align the participants head relative to the sensors of the MEG system using the *MNE-Python* module (Gramfort, 2014).

Before going into the magnetically shielded room, participants first completed a practice session with the experimenter to ensure full comprehension of the task. This consisted of 20 practice trials that were not included in the critical experiment list, but had a similar morphological structure.

MEG data were recorded continuously throughout the experiment, using a 208 channel axial gradiometer system (Kanazawa Institute of Technology, Kanazawa, Japan). A sampling rate of 1000 Hz, low-pass filter of 200 Hz and high-pass filter of .03 Hz was applied online. No band-pass filtering was applied. The experimental recording typically lasted 30 min.

### 2.5. Preprocessing

The pre-processing of MEG data was almost identical to that described in recent studies from our lab (Gwilliams et al., 2016; Gwilliams and Marantz, 2015) with the addition of ICA-based artefact removal.

First, noise-reduction was applied using the Continuously Adjusted Least Squares Method (CALM, (Adachi, 2001)) with MEG160 software (Yokohawa Electric Corporation and Eagle Technology Corporation, Tokyo, Japan), utilising eight reference channels positioned away from the participant's head. Bad channels were interpolated from surround-

ing sensors. Using the "FastICA" as implemented in *MNE-python* (Gramfort, 2014), components reflecting ocular and cardiac artefacts were removed based on visual inspection of the topographies and time course of each component. These two sources of noise are easy to detect with visual inspection. No other sources were removed. The recording was then cut into epochs – 200–600 ms around stimulus onset. Epoched sensor data were cleaned by removing any trials that exceeded a  $\pm 2000$  femto-tesla absolute or peak-to-peak threshold, which removed trials that were contaminated with movement artefacts or extraneous noise (6.9%).

The estimated location of sources from remaining trials was then computed. First, neuromagnetic data were co-registered with the FreeSurfer average brain (CorTechs Labs Inc., LaJolla, CA). The distance between the FreeSurfer and digitised fiducial points was minimised using an iterative least squares procedure in order to rotate the average brain, and then the distance between the digital head-scan and the head surface was minimised through uniform scaling, taking skull-external properties such as hair-amount into consideration.

Next, an ico-4 source space was created, containing 2562 vertices per hemisphere. For each subject a forward solution was created from the Boundary Element Model (BEM) method, and an inverse solution was computed based on the covariance matrix from the 200 ms pre-stimulus baseline period of cleaned trials. Brain activity was estimated for each vertex and time-point using dynamic statistical parametric mapping (dSPM: (Dale, 2000)). At each vertex, activity was estimated by projecting the current dipole perpendicular to the cortical surface, assigning a positive value if it is directed away from, and a negative value if it is oriented into, the cortical mass. Signed estimates were selected because they have been shown to provide superior experimental sensitivity of source-localised MEG data for the evoked response components of interest in this experiment (Gwilliams et al., 2016).

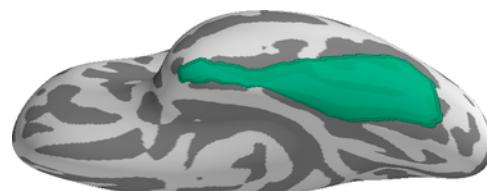
We restricted all of our analyses to the left fusiform gyrus based on the FreeSurfer aparc parcellation (Fig. 1). This choice was motivated based on previous studies that have localised M170 responses to this region, using techniques that span fMRI, MEG dipole fitting and MEG source analysis (e.g. Lewis et al., 2011; Fruchter and Marantz, 2015; Tyler, 2004; Vartiainen, 2009; Whiting et al., 2014; Hakala, 2018).

## 3. Results

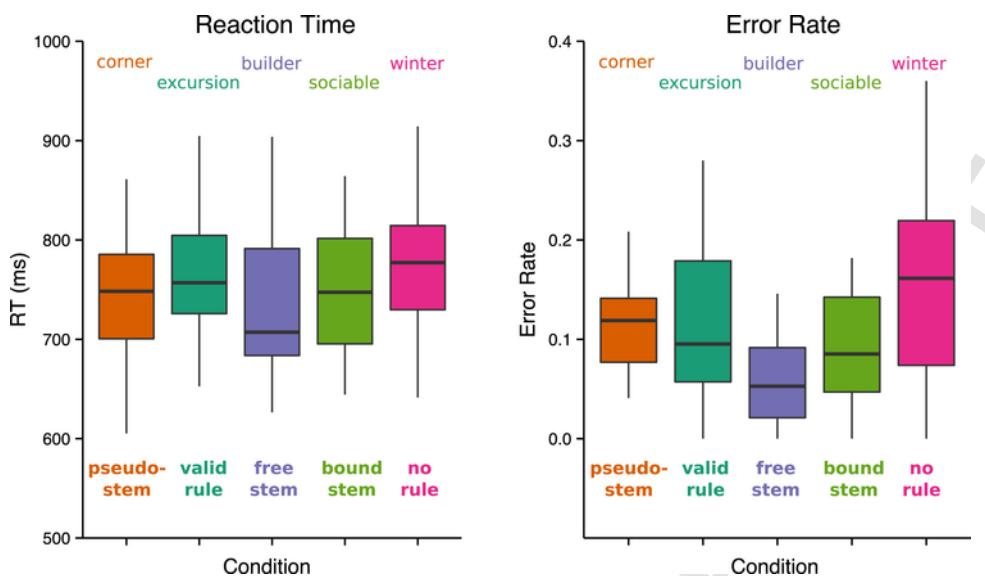
### 3.1. Grammatical wellformedness of valid word items

#### 3.1.1. Behavioural

Responses were cleaned by removing trials with reaction times  $\pm 2$  standard deviations away from the participant or item mean, which resulted in the removal of 1.9% of responses. Mean reaction time and error rates for the word items are presented in Fig. 2. For the reaction time analysis, only correct responses were included. The critical test was between the *no-rule* and *valid-rule* items: We ran a *t*-test between the two conditions and found that they behaved similarly in terms of response time ( $p = .6$ ,  $t = .52$ ) but were significantly different in accuracy ( $p = .001$ ,  $t = 3.17$ ). Indeed, *no-rule* items were incorrectly classi-



**Fig. 1.** Region of interest: Left fusiform gyrus, defined from the aparc FreeSurfer parcellation. Shown from a ventral view using an inflated surface.



**Fig. 2.** Summary of behavioural responses to word items. The thicker black line represents the median across subjects (averaged over items). The upper and lower edges of the box correspond to the first and third quartiles (the 25th and 75th percentiles). The whisker line shows the 95% confidence interval for comparing medians. Bolded names below correspond to condition labels; unbolded names above are example stimuli.

fied as non-words more often than free stem ( $p < .001$ ,  $t = 7.5$ ), bound stem ( $p < .001$ ,  $t = 4.79$ ) and pseudo-stem items ( $p = .03$ ,  $t = 2.12$ ).

### 3.1.2. MEG

**3.1.2.1. Model comparison** The goal of this analysis was to test whether the presence of a morpho-syntactic rule is sufficient to drive decomposition, in absence of an isolatable stem morpheme. In the context of our study, this boils down to whether responses to valid-rule items like *excursion* pattern with truly complex words, which we know to be treated as visually complex, or no-rule items which we know to be treated as not visually complex.

To test this, we fit a mixed effects regression model to source localised MEG responses using the *lme4* package (Bates, 2014) in R (R Core Team, 2012). In order to capture responses associated with the M170 component, we averaged activity over left fusiform gyrus, between 150 and 180 ms. This averaged neural activity served as the dependent variable. There were two critical independent variables, coding the two hypotheses that Section 1.1 outlines: i) valid-rule items are not decomposed [Hypothesis 1]; this binary variable was coded as 0 for the valid-rule and no-rule words, and as 1 for the words with an isolatable stem. ii) valid-rule items are decomposed [Hypothesis 2]; this coded the no-rule words as 0 and the valid-rule and isolatable stem words as 1. Because the items were matched on whole-word surface frequency, they necessarily differed in terms of (putative-) stem morpheme frequency. Stem frequency was therefore also included in the regression model as a control variable; setting stem frequency to be equal to whole-word frequency in the case of the no-rule and valid-rule items. For good measure, we also included the control variables that the items were matched on across items: item position, log surface frequency, log mean bigram frequency and stem length.

All variables were entered into the same mixed effects regression model, included as fixed effects and by-subject slopes. This full-model fit was then compared to a model where the fixed effect for Hypothesis 1 was removed (but remained as a by-subject slope); the same was done for Hypothesis 2. This provides an estimate of how much variance each variable was accounting for.

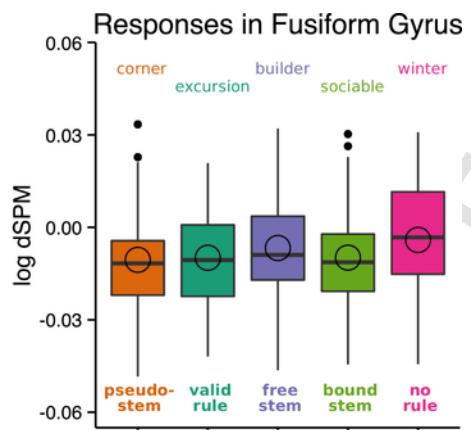
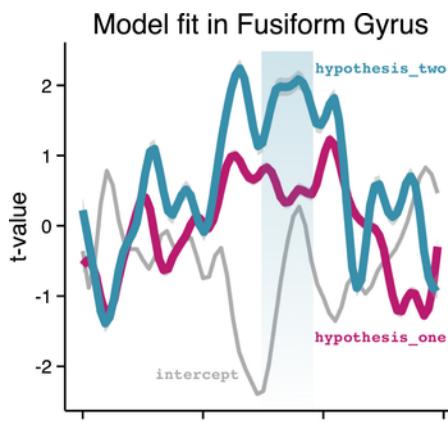
We found that Hypothesis 1 did not account for significant activation of fusiform gyrus ( $p = .66$ ,  $t = .044$ ,  $\chi = .19$ ), whereas Hypothesis 2 did ( $p = .009$ ,  $t = 2.58$ ,  $\chi = 6.69$ ). When looking at the average re-

sponse amplitude for each of the five conditions in the analysis window and region (Fig. 3, right), we see that the underlying pattern driving this response is that the no-rule items elicit a weaker amplitude than the words that contain a valid stem morpheme (note that there is weaker negative activity here, because we are using signed source estimates of the MEG data). And further, that the valid-rule words elicit a very similar response to the free-stem and pseudo-stem items.

For visualisation purposes, Fig. 3 (left) shows the coefficients of the model when fit at each time-point from 0 to 300 ms. As expected, the effect peaks around 170 ms, suggesting that this effect is indeed reflected in the M170 component.

**3.1.2.2. Transition probability** In order to push the predictions of this finding a little further, we analysed responses as a function of transition probability (TP) from stem to suffix. This variable has been found to be a reliable index of whether an item is treated by the system as having complex morphological structure. For example, there are a number of suffixes that can attach to the stem morpheme *book*, to form the complex items *book-ed*, *book-ing*, *book-s*, *book-able...*, and each suffix has a different probability of occurring with the stem. This probability, of transitioning from stem to suffix, has been shown to significantly modulate neurophysiological responses to morphologically complex words at the putative visual word form area between 150 and 200 ms, such that items with high TP elicit more activation than those with low TP (Solomyak and Marantz, 2010; Lewis et al., 2011; Simon et al., 2012).

The rationale behind our TP analysis is as follows. If valid-rule words are recognised as morphologically complex, as our first analysis suggests, then the fusiform should respond to these items as if the TP from stem to suffix is equal to 1. If, however, there is no attempt to decompose these items, the relationship between the two units is meaningless (because actually it's just one unit), and responses to these items should not pattern in accordance with this variable. In order to test this hypothesis, we averaged activity over the fusiform gyrus from 150 to 180 ms (the same spatio-temporal region analysed in the model comparison), and fit a linear model to responses to just the free stem and pseudo-stem items (i.e. *builder* and *corner* words) with TP as the critical variable as a fixed effect and by-subject slope, also including the control variables detailed above. TP significantly accounted for neural activity ( $p = .01$ ,  $t = 2.76$ ).



**Fig. 3.** Left: Timecourse of model fit in the fusiform gyrus. The y-axis represents the t-values of the model fit at each ms after word onset. Error lines around time-course traces represent standard error of the mean. Shading represents the time-window of analysis (150–180 ms). Right: Average responses to each condition in fusiform gyrus averaged in the analysis time-window. The thicker black line represents the median across subjects (averaged over items). The circles represent the mean. The upper and lower edges of the box correspond to the first and third quartiles (the 25th and 75th percentiles). The whisker line shows the 95% confidence interval for comparing medians. Individual data points are plotted when they fall outside of the confidence interval. Bolded names below correspond to condition labels; unbolded names above are example stimuli.

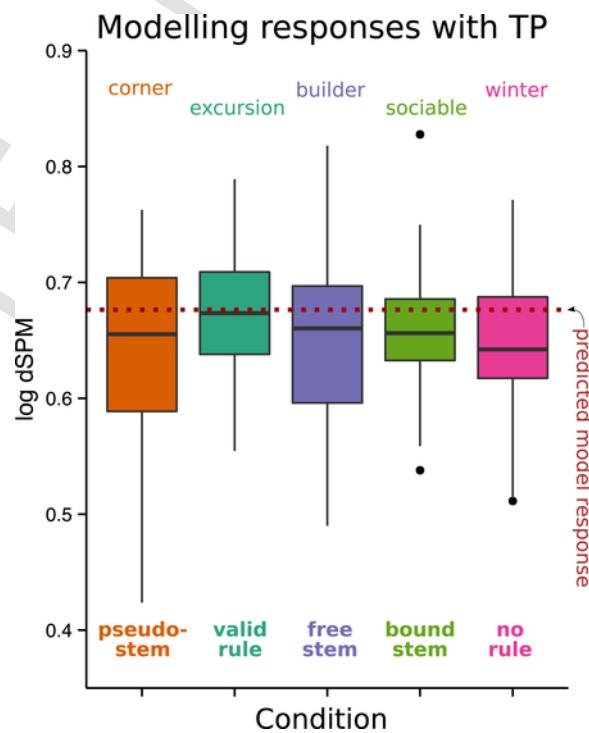
Next, we used this linear model to predict what the neural response amplitude *should* be (averaged over this region and time-window) for items that have a TP of 1. The model-predicted amplitude for items with TP = 1 (normalising for the other variables in the model like item position, log surface frequency, log mean bigram frequency and stem length) was .676 dSPM units. The critical test, then, is whether valid-rule *excursion* words elicit a response that approximates this value. If so, it suggests that TP is indeed relevant to processing these items, and therefore that these items are decomposed. Furthermore, we would expect responses to the no-rule *winter* words to be significantly different from this predicted amplitude if the no-rule items are processed as un-analysed wholes, because TP should be irrelevant to processing.

Responses to the truly complex and pseudo-complex words relative to the model-predicted amplitude are not critical here, because we have already established that responses to these words are modulated by TP — that's how we localised where and when in the cortex to run this analysis. We include these conditions in Fig. 4 for reference purposes only. But for the sake of being comprehensive, we would expect responses to be significantly less than the predicted amplitude value because they have an average transition probability that is less than 1. Again, to be clear, this doesn't mean that TP is irrelevant to processing these words; indeed, the analysis described in the previous paragraph shows that responses to free-stem and pseudo-stem items are significantly modulated by TP. We expect that the responses to the truly complex and pseudo-complex words to differ from the model-predicted amplitude of items with a TP of 1 simply because items in these conditions by definition do not have a TP of 1.

To see how closely the responses to conditions matched this predicted amplitude we ran one-sample t-tests, comparing the distribution of responses to a population mean of the expected value given the model fit. Responses to *valid-rule* items and *no-rule* items were analysed separately.

The *no-rule* items did significantly differ from the expected value ( $M = .64$ ,  $t = -3.2$ ,  $p = .001$ ) the *valid-rule* items did not significantly differ ( $M = .67$ ,  $t = -51$ ,  $p = .61$ ). This result is consistent with the interpretation that TP from stem to suffix is not relevant for the processing of *no-rule* items, but is an informative variable for the *valid-rule excursion* words. The TP analysis, then, further supports the hypothesis that *valid-rule* items are decomposed into constituent morphemes and *no-rule* items are treated as monomorphemic (see Fig. 4).

**3.1.2.3. Time-frequency analysis** A growing body of literature supports that neural responses in different frequency bands are associated with distinct neural computations (see (Bastiaansen and Hagoort, 2006a) for



**Fig. 4.** Activation in response to all conditions, averaged over the spatio-temporal cluster that was sensitive to transition probability in the free-stem words. Bolded names below correspond to condition labels; unbolded names above are example stimuli. The red dashed line represents the expected activation value given transition probability = 1. The upper and lower edges of the box correspond to the first and third quartiles (the 25th and 75th percentiles). The whisker line shows the 95% confidence interval for comparing medians. Individual data points (averaged over items) are plotted when they fall outside of the confidence intervals. Note that the *valid-rule* items pattern in accordance to what would be expected given a transition probability of 1, whereas the *no-rule* items do not.

a review of language-related studies). In studies of written-word processing, low-gamma (30–50 Hz) has been associated with communication across brain areas (Fries, 2005) memory retrieval (Pulvermüller, 1999), and responses to written words in the temporal lobe (Tanji, 2005; Crone et al., 2006; Mainy, 2007; Mainy, 2008). In lower frequencies, theta (4–7 Hz) and alpha (8–13 Hz) have been associated with semantic memory operations (Klimesch, 1997) (for a review see (Klimesch, 1999)) and working memory (Raghavachari, 2001; Rizzuto, 2006). Effects of syntactic complexity have been found in the beta

range (15–30 Hz), in comparing subject relative and object relative clauses (Weiss, 2005) as well as right-branching and centre-embedded relative clauses (Bastiaansen and Hagoort, 2006b).

As the processing stage under study is assumed to involve matching orthographic input with morpheme-form representations in memory, low-gamma is the most relevant to the computations at hand. As an additional exploratory analysis, we test how our experimental conditions differ in responses within this frequency range. In other words, we ask whether we can identify an oscillatory component associated with decompositional processes.

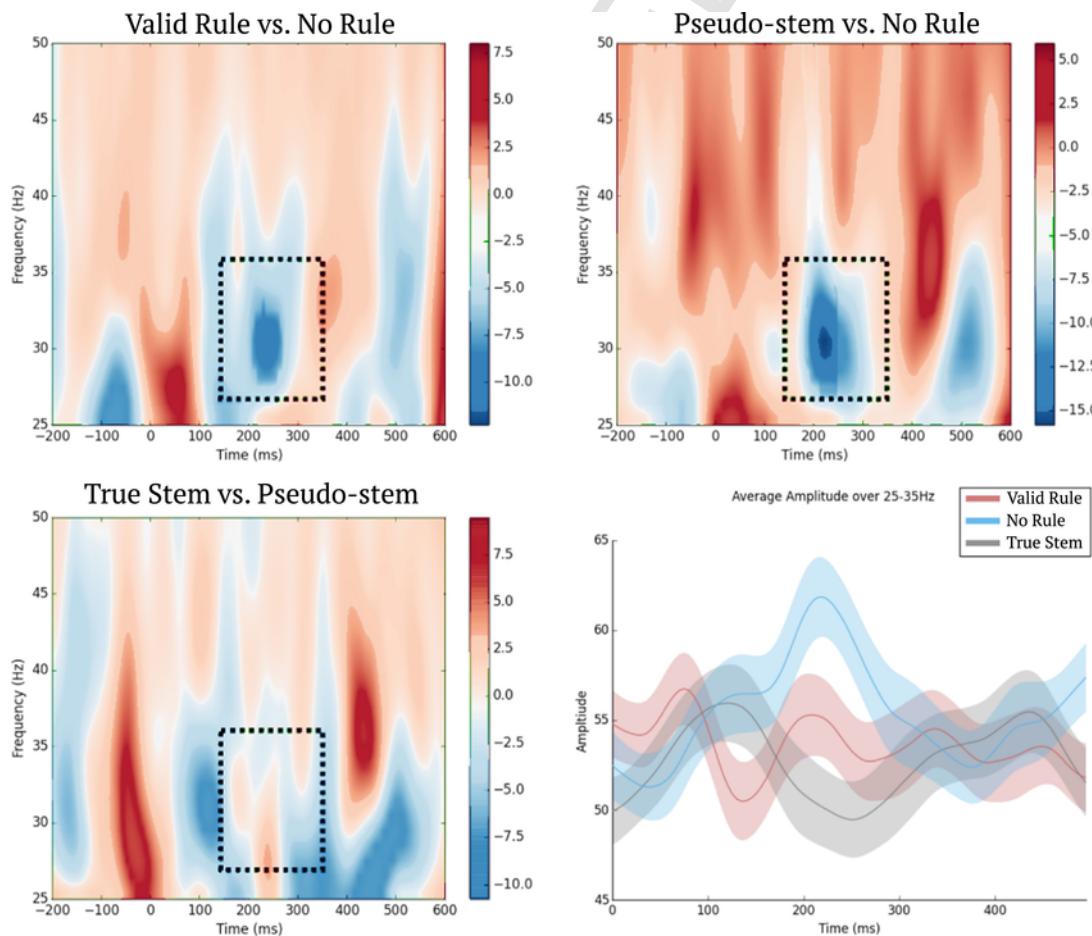
Time-frequency decomposition was applied as follows. Responses were first averaged over condition in source space; then, the time-frequency representation was computed from 25 to 50 Hz using Morlet wavelets, with 7 cycles per frequency using the *MNE-Python* module. This was applied to each source estimate of activity in the fusiform gyrus using an interval of 200 ms pre-stimulus and 600 ms post-stimulus onset.

To assess how responses differed across conditions, we used a temporal-spectral permutation cluster test as implemented in *Eelbrain* (<https://pythonhosted.org/eelbrain/>). The test was run over a 100–300 ms time-window and 25–50 Hz frequency-window. First, each data point over time and frequency, for each participant averaged over the fusiform gyrus, was used to perform a two-tailed *t*-test. This resulted in a two-dimensional matrix of un-corrected *t*-values, of the size 201 (time) x 25 (frequency). Second, clusters were formed from adjacent *t*-

values that surpassed a 1.96 threshold (equivalent to  $p = .05$ ). If the cluster was equal to or larger than 10 temporal samples and 5 spectral samples, the *t*-values in that cluster were summed to form a critical cluster statistic. Third, this value was compared to a null distribution, which was formed by randomly shuffling condition labels, performing the same *t*-test and extracting the resulting *t*-value — this random permutation was done 10,000 times. If the original test statistic fell below the 2.5th percentile or above the 97.5th percentile of this distribution of 10,000 *t*-values, the cluster was considered significant at a level of  $p < .05$ . This value was then corrected for multiple comparisons following (Maris and Oostenveld, 2007).

No-rule items (e.g., *winter*) elicited increased power in the low gamma range as compared to valid-rule items (e.g., *excursion*; 210–270 ms; 28–33 Hz,  $p = .03$ ) and pseudo-complex items (e.g., *corner*; 180–290 ms; 27–49 Hz,  $p < .01$ ). There were no differences between the two isolatable stem conditions (i.e., *corner* vs. *builder*) or the valid-rule items (summarised in Fig. 5). In other words, the only observable differences were between the no-rule items and the other conditions, suggesting that this frequency response component may reflect failure to initiate decompositional process.

This exploratory analysis corroborates our main result that valid-rule items elicit a response that is indistinguishable from words with an isolatable stem, and no-rule items elicit a significantly different response.



**Fig. 5.** Temporal-spectral cluster analysis results, showing power amplitude over time and frequency when subtracting condition responses. Top: location of cluster in the paired *t*-tests. Bottom left: no cluster found for the difference between the free stem items (dashed box indicates the absence of a difference). Bottom right: average amplitude averaged over 25–35 Hz, illustrating increased responses to the *no-rule* items. Error lines represent standard error of the mean over subjects. Note that “true stem” refers to both free and bound truly complex conditions.

### 3.2. Grammatical wellformedness of non-word items

#### 3.2.1. Behavioural

Responses to the non-word items were cleaned in the same way as the word items, resulting in removal of 1.5% of trials. Planned *t*-tests showed no difference between grammatically congruent and incongruent non-words in terms of reaction time ( $p = .44$ ,  $t = .77$ ), but there was a significant difference in accuracy in line with responses to the word items ( $p < .001$ ,  $t = 5.44$ ; see Fig. 6). This illustrates that participants were more likely to accept an item as a valid word of English when it was grammatical.

#### 3.2.2. Neurophysiological

We tested for differences between grammatically congruent and incongruent non-words in the same 150–180 ms time-window in fusiform gyrus. This was done by applying a *t*-test to these estimates of neural activity. The difference was not significant ( $p = .67$ ,  $t = .43$ ). We ran the same *t*-test in frequency space using the temporal-spectral clustering method described above, from 100 to 300 ms, between 25 and 50 Hz. No clusters were formed.

## 4. Discussion

The aim of the present study is to determine whether representations of morphemes can only be generated through exposure to the stem in multiple environments, or whether exposure to a word consistent with a morpho-syntactic rule is also capable of generating the stem. In order to address this question, we focused on two types of items: 1) *valid-rule* items such as *excursion*, which contain orthographic properties of a morphologically complex word (contain a suffix morpheme, in this case “-ion”) and obey the morpho-syntactic rule associated with the suffix (a de-verbal eventive noun), but do not have a stem that exists in the language — *excuse* is not an English verb; 2) *no-rule* items such as *winter*, which again have orthographic properties matching that of a morphologically complex word, but do not adhere to the grammatical behaviour required by that suffix. Both kinds of items are considered monomorphic from the standpoint of exposure (and annotated as such in the ELP and CELEX), because their potential stems do not occur in any other word or in isolation; however, if mor-

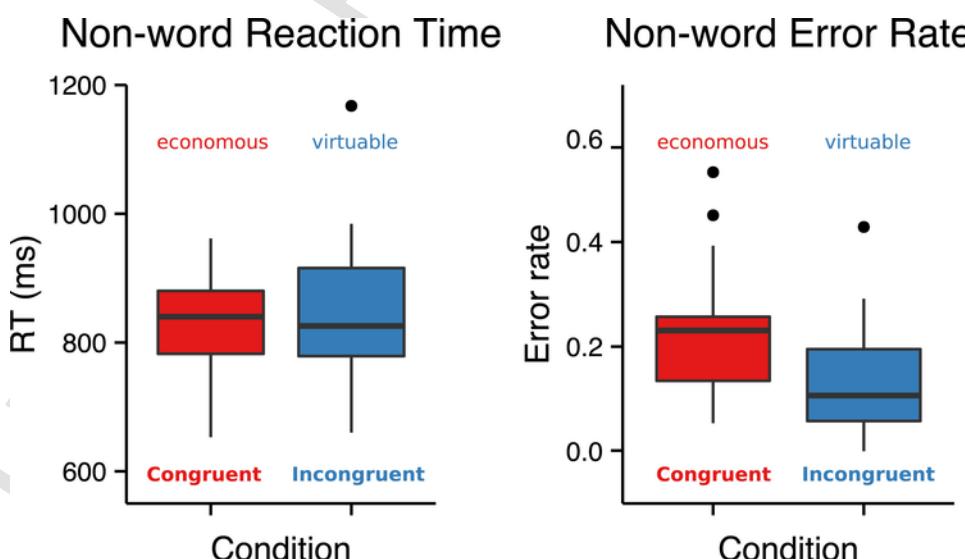
phological representations can be formed by applying valid grammatical rules, words like *excursion* should be decomposed similar to words like *builder*.

Our results confirm that exposure to words that obey morpho-syntactic rules leads to the formation of constituent morphological representations. They illustrate differential processing of items that vary only in whether a rule is present, suggesting that while items like *excursion* lead to access to an abstract visual form representation of the unique stem “excuse”, items with identical surface structure but no grammatical rule like *winter* do not lead to the representation of “wint”.

#### 4.1. Decomposition in fusiform gyrus

In our first analysis, we modelled responses in the fusiform gyrus, either assuming i) only items with an isolatable stem will be considered visually complex or ii) items with *either* an isolatable stem or a valid morpho-syntactic rule will be considered complex. The second model best accounted for neural responses. Given the wealth of prior literature suggesting that free and bound stem words are decomposed, this boils down to the result that while *no-rule* words are not decomposed (they are the odd ones out, replicating previous studies) *valid-rule* words are decomposed. This neurophysiological result was also mirrored in the behavioural error rates: Participants made significantly more errors by rejecting *no-rule* items as valid English words as compared to all other conditions. Notably, the reaction time data did not follow the same trend, and *valid-rule* items were responded to significantly slower than free and bound stem words. We attribute this slow-down to the fact that *valid-rule* words were matched on surface frequency but not stem frequency to the truly complex items. Stem frequency has been found to correlate with reaction time in similar visual lexical decision studies, such that words with lower frequency stems take longer to respond to (Colé et al., 1997); this is perfectly in line with our observation.

This finding suggests, therefore, that morphological representations can be formed through the exploitation of morpho-syntactic rules. From this perspective, what is important is not repeated exposure to the morphological constituents per se (which is null in the *excursion* case, but we find it is decomposed regardless). Rather, it is exposure to a word that obeys the morpho-syntactic rule of its suffix. One could imagine that these rules are learnt from repeated association between a



**Fig. 6.** Summary of behavioural responses to non-word conditions. The thicker black line represents the median across subjects (averaged over items). The upper and lower edges of the box correspond to the first and third quartiles (the 25th and 75th percentiles). The whisker line shows the 95% confidence interval for comparing medians. Individual data points are plotted when they fall outside of the confidence interval.

suffix (e.g. -ion) and its function (form an eventive noun, derived from a verb stem). And, when a word is encountered that follows the learnt syntactic rule associated with the suffix, a representation is formed of the stem morpheme. Subsequent exposure to this word allows it to be recognised via the representation that was formed.

Why then, would pseudo-complex words like *corner* and *brother* be considered visually complex, given that they do not follow the grammatical rule? It seems that the parser may be sensitive to the presence of a visual morpheme-form, regardless of the rule that was applied to create it. From this perspective, it is previous encounters with the stem *corn* being used in a grammatically valid way (e.g. corn + Ø, corny) that generates the abstract representation of *corn*, which is later recognised within the word *corner*. Finding indistinguishable responses to valid rule and pseudo-stem words therefore suggests that the driving force of decomposition is not the semantic relationship between constituents, but rather recognition of an abstract stored representation of those constituents. When the parser recognises the visual morpheme-forms for *corn* and -er, the word is decomposed; it would not be until later that the system recognises that the stem and the whole word are not related.

The utility of morpho-syntactic rules is particularly prevalent when there is not a one-to-one mapping between the constituents in the complex form and their visual presentation in isolation. A related study comes from (McCormick et al., 2008): The authors tested whether, and the extent to which, the morphological parser is robust to the orthographic alternations often found in complex words. For example, dropping the “e” in *adore* → *adorable*, or duplicating a consonant in *drop* → *dropper*. The authors found that morphologically complex words were treated the same by the language system, regardless of the presence of these orthographic alternations. This finding suggests that the brain can recognise that an item is complex, even with imperfect orthographic overlap.

Relatedly, Crepaldi et al. (2010) found that words with irregular inflections fell → fall also appear to be processed as morphologically complex, even though these irregular forms do not follow a systematic orthographic transformation. Our results suggest that while this may be true, the system is sensitive to the difference between free stem and bound stem items at the M170, possibly indicating that the degree of orthographic overlap between the stem and the complex whole still has a role to play (see Fig. 3, right). These previous results, combined with our own, suggest that complexity is determined based on the applicability of abstract rules rather than lower-level visual features.

Pulling everything together, while our results suggest that morpho-syntactic rules are used to generate representations, they do not speak to the nature of how the representations themselves are coded. However, when we couple these findings with previous results, it seems that they need to be abstract enough to overcome orthographic inconsistencies, so are unlikely to closely resemble the written input. Further work needs to be conducted in order to more precisely understand the nature of these morphological representations, and how exactly they are encoded in the brain.

#### 4.2. Transition probability between morphemes

As an additional test we analysed activation as a function of TP from stem to suffix. In previous studies (Solomyak and Marantz, 2010; Lewis et al., 2011; Simon et al., 2012) this variable significantly modulated responses to morphologically complex words, indexing morphological decomposition. If valid-rule items are decomposed, we would expect responses to these items to pattern as if they have a TP of 1 (because there is no other suffix that attaches to the stem). Recall that the transition probability variable refers specifically to the relation between morphological constituents, not between letter strings.

If no decomposition occurs, then this variable should be irrelevant to processing and responses should not pattern systematically. In corroboration with our previous analyses, we indeed find that responses to valid-rule items scale with TP — they act as if they have a TP of 1. Importantly, the no-rule items did not systematically pattern with TP, also in agreement with the previous results suggesting that they are not decomposed.

This finding is particularly important because for items with TP close to 1, access to the stem does not grant significantly more information than access to the whole word. For example, TP of the pseudo complex word *broth-er* is very close to 1, because the stem *broth* occurs much more often with the suffix “-er” than without. This is necessarily true for all of the excursion-type items as well, because the frequency of the complex item is the frequency of the stem. These results can only be accounted for by a theory that assumes decomposition into morphemes as initiated by the morpho-syntactic rules that govern constituents. And importantly suggests that variables reflecting exposure, such as frequency of constituents and whole forms, serve to moderate rather than dictate the decompositional process.

#### 4.3. Oscillatory components of decomposition

In order to better understand the neurophysiological responses underlying this result, we analysed activity over time and frequency in the fusiform gyrus. No-rule words like *winter* elicited greater power in the low-gamma range (~25–35 Hz) at 200 ms after word onset, in comparison to both valid-rule items (e.g. *excursion*) and pseudo-stem items (e.g. *corner*). This frequency component was not observable in the difference between the two isolatable stem conditions (e.g. *corner* vs. *builder*), the valid-rule and the free stem items (e.g. *excursion* vs. *builder*) or the non-words containing stems (e.g. *drinkage* vs. *finalic*), suggesting that the locus of this effect is the comparison between items with and without constituent morpheme-form representations.

Observing such responses in the low-gamma range is consistent with previous research, which has associated this frequency band with the integration and binding of information, as well as reflecting access to abstract lexical information (Bastiaansen and Hagoort, 2006a; Mainy, 2007). Specifically within the domain of language comprehension, differential responses have been reported ~30 Hz for the comparison between mono-morphemic words and pseudo-words between 100 and 250 ms after onset (Lutzenberger et al., 1994; Pulvermüller, 2006) word classes (Pulvermüller, 1996) and semantic differences (see Pulvermüller, 2005; Pulvermüller et al., 2006 for a review). Because there was no difference between the processing of grammatically congruent and incongruent non-words within this time-window and frequency range, this component seems not to be simply sensitive to the congruity of our stimulus items. Instead, it appears to reflect an attempt to access a morpheme-form representation and match it to the input: Decomposition is successful for words like *excursion*, *corner* and *builder* because they can be exhaustively parsed into stem + affix; decomposition fails for *winter* items, however, because there is no entry for “wint” once the pseudo-suffix is removed — requiring recognition of the full form *winter* instead. We interpret increased low-gamma responses to no-rule items as reflecting a “failure to bind” the non-stem with a morphological representation; though, future work will need to be conducted in order to properly substantiate this claim.

#### 5. Conclusion

By modelling activity in the fusiform gyrus in response to written words, we tested which kinds of words are considered visually complex by the processing system, in order to deduce how the brain generates morphological representations. Our results suggest that a lexical item is

considered visually complex if it obeys the morpho-syntactic rule associated with its suffix. Concretely, we find that valid-rule words like *excursion* pattern alongside uncontroversially decomposed words like *builder* and *sociable*, whereas the no-rule words like *winter* elicit a separable response and are not decomposed. Any morphologically complex word that obeys the grammatical rules of the language seems to be represented via its constituent morphemes - even if the morphemes have never been encountered in isolation - and it is these constituents that support subsequent recognition and processing.

### Acknowledgements

This work was supported by the NYU Abu Dhabi Institute under Grant G1001 (AM). We would like to thank Davide Crepaldi for his

helpful comments on an earlier version of the manuscript, as well as Linnaea Stockall for interesting and useful discussion when designing the study.

### Appendix A

See Table A1

**Table A1**  
List of stimulus items.

Bound stem	Free Stem	Pseudo-complex	Valid Rule	No Rule
Adorable	Agreeable	Portable	Affable	Parable
Arguable	Bearable	Bandage	Amiable	Beverage
Curable	Passable	Footage	Culpable	Montage
Likable	Suitable	Hostage	Palpable	Mortgage
Lovable	Workable	Massage	Carnage	Sausage
Memorable	Bondage	Pillage	Envisage	Vantage
Movable	Leakage	Plumage	Espionage	Village
Notable	Leverage	Rampage	Foliage	Pedestal
Sizable	Llinkage	Teenage	Garbage	Elephant
Sociable	Mileage	Vintage	Heritage	Enchant
Valuable	Package	Rational	Salvage	Pheasant
Barrage	Postage	Sergeant	Suffrage	Boulevard
Storage	Shortage	Bustard	Abysmal	Custard
Approval	Orbital	Buzzard	Admiral	Haggard
Arsenal	Voltage	Mustard	Genital	Leopard
Decimal	Cynical	Boarder	Initial	Orchard
Familial	Frontal	Bouncer	Marital	Frigate
Glacial	lLogical	Brother	Neutral	Magnate
Imperial	Magical	Burgher	Nominal	Template
Maximal	Optical	Cracker	Radical	Blubber
Minimal	Topical	Dresser	Thermal	Bolster
Musical	Boundary	Flicker	Vertical	Boulder
Natural	Dietary	Knocker	Merchant	Brazier
Stoppage	Admirer	Pitcher	Valiant	Clatter
Sensual	Adviser	Sleeper	Verdant	Clobber
Spatial	Bloomer	Sneaker	Bastard	Cylinder
Buoyant	Booster	Splinter	Acetate	Fritter
Defiant	Breaker	Sticker	Agitate	Glacier
Radiant	Breeder	Sweater	Detonate	Glimmer
Variant	Builder	Trainer	Dictate	Juniper
Pulsate	Catcher	Trooper	Elevate	Lacquer
Urinate	Charger	Voucher	Emulate	Lobster
Chaotic	Charmer	Blanket	Imitate	Platter
Idyllic	Cleaner	Brisket	Isolate	Prosper
Melodic	Creeper	Cabinet	Migrate	Shudder
Pacific	Cruiser	Cricket	Narrate	Slobber
Prosaic	Doubter	Hatchet	Violate	Slumber
Diction	Dreamer	Thicket	Butcher	Smother
Elation	Drifter	Hackney	Cadaver	Snicker
Erosion	Drinker	Trolley	Holster	Stammer
Suction	Fiddler	Electric	Monster	Terrier
Tension	Fielder	Organic	Trigger	Banquet
Brevity	Fighter	Faction	Bayonet	Cabaret
Clarity	Floater	Charity	Bouquet	Crochet
Density	founder	Priority	Bracket	Croquet
Trinity	Gambler	Gauntlet	Epithet	Parapet
Anxious	Hustler	Leaflet	Ratchet	Chimney
Envious	Llighter	Scarlet	Boloney	Journey
Fibrous	Marcher	Callous	Paisley	Caution
Furious	Mariner	Ravenous	Crucial	Cushion
Piteous	Painter	Riotous	Chronic	Fiction
Vacuous	Teacher	Vicarious	Drastic	Oblivion
Zealous	Twister	Classic	Serious	Awkward

## References

- Adachi, Y., et al., 2001. Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. *IEEE Trans. Appl. Supercond.* 11 (1), 669–672.
- Baayen R.H. et al., 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychol. Rev.* 118 (3), 438.
- Baayen, R.H., Piepenbrock, R., van H, R., 1993. The {CELEX} lexical data base on {CD-ROM}..
- Balota D.A. et al., 2007. The English lexicon project. *Behav. Res. Methods* 39 (3), 445–459.
- Bastiaansen, M., Hagoort, P., 2006a. Oscillatory neuronal dynamics during language comprehension 159, pp. 179–196.
- Bastiaansen, M., Hagoort P. 2006. Oscillatory neuronal dynamics during language comprehension. *Prog. Brain Res.* 159, 179–196.
- Bates, D., et al., 2014. Fitting linear mixed-effects models using lme4. arXiv preprint arXiv: 1406.5823.
- Caramazza, A., Laudanna, A., Romani, C., 1988. Lexical access and inflectional morphology. *Cognition* 28 (3), 297–332.
- Cavalli, E., et al., 2016. Spatiotemporal dynamics of morphological processing in visual word recognition. *J. Cogn. Neurosci.* (2016).
- Clahsen, H., 2003. Derivalional morphology in the German mental lexicon: a dual mechanism account. In: Clahsen, Harald, Sonnenstuhl, Ingrid, Blevins, James P. (Eds.) *Morphological Structure in Language Processing*. 151, p. 125.
- Cohen, L., et al., 2000. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123 (2), 291–307.
- Colé P. Beauvillain C. Segui J. 1989. On the representation and processing of prefixed and suffixed derived words: a differential frequency effect. *J. Mem. Lang.* 28 (1), 1–13.
- Colé P. Segui J. Taft M. 1997. Words and morphemes as units for lexical access. *J. Mem. Lang.* 37 (3), 312–330.
- Crepaldi D. et al., 2010. ‘Fell’primes ‘fall’, but does ‘bell’prime ‘ball’? Masked priming with irregularly-inflected primes. *J. Mem. Lang.* 63 (1), 83–99.
- Crone N.E. Sinai A. Korzeniewska A. 2006. High-frequency gamma oscillations and human brain mapping with electrocorticography. *Prog. Brain Res.* 159 275–295.
- Dale A.M. et al., 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron* 26 (1), 55–67.
- Dehaene, S., Naccache, L., 2001. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79 (1), 1–37.
- Deutsch, A., Frost R. Forster K.I. 1998. Verbs and nouns are organized and accessed differently in the mental lexicon: evidence from Hebrew. *J. Exp. Psychol.: Learn. Mem. Cogn.* 24 (5), 1238.
- Devlin, J.T., et al., 2004. Morphology and the internal structure of words. *Proc. Natl. Acad. Sci. USA* 101 (41), 14984–14988.
- Diependael K. Sandra D. Grainger J. 2009. Semantic transparency and masked morphological priming: the case of prefixed words. *Mem. Cogn.* 37 (6), 895–908.
- Dufaubeitia J.A. Perea M. Carreiras M. 2007. Do transposed-letter similarity effects occur at a morpheme level? Evidence for morpho-orthographic decomposition. *Cognition* 105 (3), 691–703.
- Feldman L.B. et al., 2004. What do graded effects of semantic transparency reveal about morphological processing? *Brain Lang.* 90 (1–3), 17–30.
- Feldman, L.B., O'Connor, P.A., del Prado Martín, F.M., 2009. Early morphological processing is morphosemantic and not simply morpho-orthographic: a violation of form-then-meaning accounts of word recognition. *Psychon. Bull. Rev.* 16 (4), 684–691.
- Fries P. 2005. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9 (10), 474–480.
- Fruchter J. Marantz A. 2015. Decomposition, lookup, and recombination: meg evidence for the full decomposition model of complex visual word recognition. *Brain Lang.* 143 81–96.
- Fruchter J. Stockall L. Marantz A. 2013. MEG masked priming evidence for form-based decomposition of irregular verbs. *Front. Hum. Neurosci.* 7.
- Giraudo H. Grainger J. 2000. Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Lang. Cogn. Process.* 15 (4–5), 421–444.
- Gonneman L.M. Seidenberg M.S. Andersen E.S. 2007. Graded semantic and phonological similarity effects in priming: evidence for a distributed connectionist approach to morphology. *J. Exp. Psychol. Gen.* 136 (2), 323–345.
- Grainger J. Colé P. Segui J. 1991. Masked morphological priming in visual word recognition. *J. Mem. Lang.* 30 (3), 370–384.
- Gramfort A. et al., 2014. MNE software for processing MEG and EEG data. *Neuroimage* 86 446–460.
- Gwilliams L. Marantz A. 2015. Non-linear processing of a linear speech stream: the influence of morphological structure on the recognition of spoken Arabic words. *Brain Lang.* 147 1–13.
- Gwilliams, L., Lewis, G., Marantz, A., 2016. Functional characterisation of letter-specific responses in time, space and current polarity using magnetoencephalography. *NeuroImage* 132, 320–333.
- Gwilliams, L.E., Monahan, P.J., Samuel A.G. 2015. Sensitivity to morphological composition in spoken word recognition: Evidence from grammatical and lexical identification tasks. *J. Exp. Psychol.: Learn., Mem. Cogn.* 41, (6), 1663.
- Hakala, T. Hakala, et al., Information properties of morphologically complex words modulate brain activity during word reading, *Hum. Brain Mapp.* (2018).
- Klimesch, W. 1997. W. Klimesch, EEG-alpha rhythms and memory processes. *Int. J. Psychophysiol.* 26 (1) (1997) 319–340.
- Klimesch, W., 1999. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* 29 (2), 169–195.
- Lavric, A., Clapp A. Rastle K. 2007. ERP evidence of morphological analysis from orthography: a masked priming study. *J. Cogn. Neurosci.* 19 (5), 866–877.
- Lehtonen M. Monahan P.J. Poeppel D. 2011. Evidence for early morphological decomposition: combining masked priming with magnetoencephalography. *J. Cogn. Neurosci.* 23 (11), 3366–3379.
- Lewis G. Solomyak O. Marantz A. 2011. The neural basis of obligatory decomposition of suffixed words. *Brain Lang.* 118 (3), 118–127.
- Longtin, C.-M., Meunier F. 2005. Morphological decomposition in early visual word processing. *J. Mem. Lang.* 53 (1), 26–41.
- Lutzenberger, W., Pulvermüller F. Birbaumer N. 1994. Words and pseudowords elicit distinct patterns of 30-Hz EEG responses in humans. *Neurosci. Lett.* 176 (1), 115–118.
- Mainy N., et al., 2007. Neural correlates of consolidation in working memory. *Hum. Brain Mapp.* 28, (3), 183–193.
- Mainy N. et al., 2008. Cortical dynamics of word recognition. *Hum. Brain Mapp.* 29 (11), 1215–1230.
- Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164 (1), 177–190.
- Marslen-Wilson, W., et al. 1994. Morphology and meaning in the English mental lexicon. *Psychol. Rev.* 101 (1), 3.
- Mccormick, S.F., Rastle, K., Davis, M.H., 2008. Is there a ‘fete’ in ‘fetish’? Effects of orthographic opacity on morpho-orthographic segmentation in visual word recognition. *J. Mem. Lang.* 58 (2), 307–326.
- Meunier, F., Longtin, C.-M., 2007. Morphological decomposition and semantic integration in word processing. *J. Mem. Lang.* 56 (4), 457–471.
- Morris J. Grainger J. Holcomb P.J. 2008. An electrophysiological investigation of early effects of masked morphological priming. *Lang. Cogn. Process.* 23 (7), 1021–1056.
- Morris J. Grainger J. Holcomb P.J. 2013. Tracking the consequences of morpho-orthographic decomposition using ERPs. *Brain Res.* 1529 92–104.
- Pastizzo M.J. Feldman L.B. 2002. Discrepancies between orthographic and unrelated baselines in masked priming undermine a decompositional account of morphological facilitation. *J. Exp. Psychol.: Learn. Mem. Cogn.* 28 (1), 244–249.
- Pinker S. Prince A. 1988. On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition* 28 (1–2), 73–193.
- Pulvermüller, F., et al., 2006. Tracking speech comprehension in space and time. *NeuroImage* 31 (3), 1297–1305.
- Pulvermüller, F., et al., 1996. Brain rhythms of language: nouns versus verbs. *Eur. J. Neurosci.* 8 (5), 937–941.
- Pulvermüller F. 1999. Words in the brain’s language. *Behav. Brain Sci.* 22 (2), 253–279.
- Pulvermüller, F., 2005. Brain mechanisms linking language and action. *Nat. Rev. Neurosci.* 6 (7), 576–582.
- Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R., 2006. Brain signatures of meaning access in action word recognition. *Brain* 17 (6).
- Pylkkänen L. et al., 2004. Neural correlates of the effects of morphological family frequency and family size: an MEG study. *Cognition* 91 (3), B35–B45.
- Pylkkänen L. Marantz A. 2003. Tracking the time course of word recognition with MEG. *Trends Cogn. Sci.* 7 (5), 187–189.
- Raghavachari S., et al., 2001. Gating of human theta oscillations by a working memory task. *J. Neurosci.* 21 (9), 3175–3183.
- Rastle K. et al., 2000. Morphological and semantic effects in visual word recognition: a time-course study. *Lang. Cogn. Process.* 15 (4–5), 507–537.
- Rastle K. Davis M.H. 2008. Morphological decomposition based on the analysis of orthography. *Lang. Cogn. Process.* 23 (7–8), 942–971.
- Rastle K. Davis M.H. New B. 2004. The broth in my brother’s brothel: morpho-orthographic segmentation in visual word recognition. *Psychon. Bull. Rev.* 11 (6), 1090–1098.
- Rizzuto D.S. et al., 2006. Human neocortical oscillations exhibit theta phase differences between encoding and retrieval. *NeuroImage* 31 (3), 1352–1358.
- Seidenberg M.S. Gonnerman L.M. 2000. Explaining derivational morphology as the convergence of codes. *Trends Cogn. Sci.* 4 (9), 353–361.
- Simon D.A. Lewis G. Marantz A. 2012. Disambiguating form and lexical frequency effects in MEG responses using homonyms. *Lang. Cogn. Process.* 27 (2), 275–287.
- Solomyak O. Marantz A. 2010. Evidence for early morphological decomposition in visual word recognition. *J. Cogn. Neurosci.* 22 (9), 2042–2057.
- Taft M. 1979. Recognition of affixed words and the word frequency effect. *Mem. Cogn.* 7 (4), 263–272.
- Taft M. Forster K.I. 1975. Lexical storage and retrieval of prefixed words. *J. Verbal Learn. Verbal Behav.* 14 (6), 638–647.
- Tanji K. et al., 2005. High-frequency γ-band activity in the basal temporal cortex during picture-naming and lexical-decision tasks. *J. Neurosci.* 25 (13), 3287–3293.
- Tyler L. 2004. Neural processing of nouns and verbs: the role of inflectional morphology. *Neuropsychologia* 42 (4), 512–523.
- Vartiainen J. et al., 2009. Neural dynamics of reading morphologically complex words. *NeuroImage* 47 (4), 2064–2072.
- Weiss S. et al., 2005. Increased neuronal communication accompanying sentence comprehension. *Int. J. Psychophysiol.* 57 (2), 129–141.
- Whiting C. Shtyrov Y. Marslen-Wilson W. 2014. Real-time functional architecture of visual word recognition. *J. Cogn. Neurosci.*
- Whiting C. Shtyrov Y. Marslen-Wilson W. 2015. Real-time functional architecture of visual word recognition. *J. Cogn. Neurosci.* 27 (2), 246–265.
- Zweig E. Pylkkänen L. 2009. A visual M170 effect of morphological complexity. *Lang. Cogn. Process.* 24 (3), 412–439.