CORTEX 106 (2018) 213–236

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

Building words and phrases in the left temporal lobe

Graham Flick [a,*,1], Yohei Oseki [b,**,1], Amanda R. Kaczmarek [c],
Meera Al Kaabi [d], Alec Marantz [a,b,e] and Liina Pylkkänen [a,b,e]

[a] NYUAD Institute, New York University Abu Dhabi, United Arab Emirates
[b] Department of Linguistics, New York University, USA
[c] Department of Psychological and Brain Sciences, University of California, Santa Barbara, USA
[d] Department of Linguistics, United Arab Emirates University, United Arab Emirates
[e] Department of Psychology, New York University, USA

ARTICLE INFO

Article history:
Received 19 December 2017
Reviewed 13 January 2018
Revised 3 April 2018
Accepted 6 June 2018
Action editor Alessandro Tavano
Published online 21 June 2018

Keywords:
Semantic composition
Magnetoencephalography
Anterior temporal cortex
Complex word recognition

ABSTRACT

A central part of knowing a language is the ability to combine basic linguistic units to form complex representations. While our neurobiological understanding of how words combine into larger structures has significantly advanced in recent years, the combinatory operations that build words themselves remain unknown. Are complex words such as *tombstone* and *starlet* built with the same mechanisms that construct phrases from words, such as *grey stone* or *bright star*? Here we addressed this with two magnetoencephalography (MEG) experiments, which simultaneously varied demands associated with phrasal composition, and the processing of morphological complexity in compound and suffixed nouns. Replicating previous findings, we show that portions of the left anterior temporal lobe (LATL) are engaged in the combination of modifiers and monomorphemic nouns in phrases (e.g., *brown rabbit*). As regards compounding, we show that semantically transparent compounds (e.g., *tombstone*) also engage left anterior temporal cortex, though the spatiotemporal details of this effect differed from phrasal composition. Further, when a phrase was constructed from a modifier and a transparent compound (e.g., *granite tombstone*), the typical LATL phrasal composition response appeared at a delayed latency, which follows if an initial within-word operation (*tomb* + *stone*) must take place before the combination of the compound with the preceding modifier (*granite* + *tombstone*). In contrast to compounding, suffixation (i.e., *star* + *let*) did not engage the LATL in any consistent way, suggesting a distinct processing route. Finally, our results suggest an intriguing generalization that morpho-orthographic complexity that does not recruit the LATL may block the engagement of the LATL in subsequent phrase building. In sum, our findings offer a detailed spatiotemporal characterization of the lowest level combinatory operations that ultimately feed the composition of full sentences.

© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

---

* *Corresponding author.* NYUAD Institute, New York University Abu Dhabi, United Arab Emirates.
** *Corresponding author.* Department of Linguistics, New York University, USA
E-mail addresses: graham.flick@nyu.edu (G. Flick), yohei.oseki@nyu.edu (Y. Oseki).

[1] denotes equal contribution.

https://doi.org/10.1016/j.cortex.2018.06.004

0010-9452/© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Human language is a generative system allowing for the composition of infinitely complex expressions from smaller parts. The smallest meaningful units of language are morphemes, the elementary pieces that make up complex words (e.g., *neuro-scien-tist*). Complex words can then enter into further combinatory operations to form phrases (e.g., *an excit-ed neuro-scien-tist*), and such phrases can combine with each other to form sentences (e.g., *an excit-ed neuro-scien-tist publish-ed a comment-ary*), which become the building blocks of discourse. A fundamental question concerning the neural bases of language is the extent to which combinatory operations at these different levels are computationally uniform or distinct, with relevant theoretical accounts differing sharply in their proposals concerning the similarities and differences (see Halle & Marantz, 1993; Marantz, 1997; c.f., Aronoff, 1994). Recently, there has been great progress towards understanding the mechanisms that underlie composition between words, with various lines of research identifying brain responses related to the construction of sentence meaning (Fedorenko et al., 2016; Frankland & Greene, 2015) and hierarchical linguistic structures (Ding, Melloni, Zhang, Tian, & Poeppel, 2016; Nelson et al., 2017), as well as functional subprocesses that take place during the composition of minimal phrases (Bemis & Pylkkänen, 2011; Blanco-Elorrieta & Pylkkänen, 2016; Del Prato & Pylkkänen, 2014; Poortman & Pylkkänen, 2016; Westerlund & Pylkkänen, 2014; Zhang & Pylkkänen, 2015; Ziegler & Pylkkänen, 2016). Relatively unstudied however, is the question of how the neural processes that underlie potential instances of composition within words relate to those associated with composition between words.

Here, we report the results of a magnetoencephalography (MEG) investigation that tested for the presence of a general combinatory operation supporting not only phrasal composition, taking place between pairs of words, but also two varieties of morphological composition, taking place between pairs of morphemes. To do so, we first performed a functional localizer-like analysis to identify left hemisphere regions that showed sensitivity to modifier-noun composition involving monomorphemic nouns. By then adopting these areas as functional regions of interest (fROIs) and varying phrasal composition and word-internal morphological complexity in the same design, we were able to test whether those regions associated with phrasal composition were similarly sensitive to composition within words. Additionally, by further modifying nouns containing morpho-orthographic complexity, we were also able to investigate how the typical LATL combinatory response may be modulated by various types of complexity present in the nouns that enter into composition.

1.1. *Prior work on phrasal composition*

Within the extant literature, it is commonly proposed that a large number of broadly defined brain areas may each play a role, or multiple roles, in the neurobiological underpinnings of composition, with functional sub-divisions possibly existing at fine-grained spatial demarcations (for relevant reviews see Friederici, 2012; Hagoort and Indefrey, 2014). These areas typically include the left inferior frontal gyrus, anterior, middle, and posterior sections of the left temporal lobe, and cortex surrounding the left temporo-parietal junction, including the angular gyrus. Within the work on this topic, and across this network of regions, one of the most consistent generalizations to date is the involvement of the left anterior temporal lobe (LATL); a finding that has been reported across studies focused on sentential (Brennan & Pylkkänen, 2012; Friederici, Opitz, & Cramon, 2000; Humphries, Binder, Medler, & Liebenthal, 2006; Humphries, Love, Swinney, & Hickok, 2005; Mazoyer et al., 1993; Pallier, Devauchelle, & Dehaene, 2011; Rogalsky & Hickok, 2009; Stowe et al., 1998; Vandenberghe, Nobre, & Price, 2002; Xu, Kemeny, Park, Grattali, & Braun, 2005) and phrasal (Bemis & Pylkkänen, 2011, 2012, 2013a, 2013b; Blanco-Elorrieta & Pylkkänen, 2016; Bozic, Fontenaeu, Su, & Marslen-Wilson, 2015; Del Prato & Pylkkänen, 2014; Westerlund & Pylkkänen, 2014; Westerlund, Kastner, Al Kaabi, & Pylkkänen, 2015; Zhang & Pylkkänen, 2015; Ziegler & Pylkkänen, 2016) composition. A subset of the latter studies, which has used MEG to elucidate spatiotemporal correlates of phrasal composition, has provided a temporal profile of the LATL's involvement in minimal modifier-noun phrases (e.g., *red boat*). The results of this work have consistently demonstrated an increased LATL response to nouns preceded by a modifier, relative to non-combinatory control conditions (e.g., a list of two words), beginning 150–250 ms after the onset of the noun (Bemis & Pylkkänen, 2011, 2012, 2013a, 2013b; Westerlund et al., 2014).

Following the initial identification of this LATL-localizing MEG effect, more recent work has begun to delineate the precise computational role of this response, suggesting that the increased activity may index a stage at which semantic or conceptual representations are combined. This characterization comes from results showing that the typical 150–250 ms LATL response is sensitive to the conceptual specificity of both the modifier and the noun (Westerlund & Pylkkänen, 2014; Zhang & Pylkkänen, 2015; Ziegler & Pylkkänen, 2016), and that, at least during production, it may only be elicited by cases of conceptual combination, with instances of numeral quantification failing to drive it (Del Prato & Pylkkänen, 2014; Blanco-Elorrieta & Pylkkänen, 2016; for related results and proposals see; Patterson, Nestor, & Rogers, 2007; Baron & Osherson, 2011; Chadwick et al., 2016; Boylan, Trueswell, & Thompson-Schill, 2017). Although, on occasion, MEG studies employing the minimal modifier-noun paradigm have found accompanying increases in other regions, such as the left angular gyrus (Bemis & Pylkkänen, 2012) and ventromedial prefrontal cortex (Bemis & Pylkkänen, 2011; Pylkkänen, Bemis, & Elorrieta, 2014), these effects have been inconsistent across studies. While the lesser consistency of vmPFC findings as compared to the LATL is likely due to the relative noisiness of activity localized to the prefrontal regions (being close to the eyes), there is recent evidence suggesting that MEG activity localizing in regions near to the angular gyrus may reflect the relationality of concepts, instead of the process of composition (Williams, Reddigari, & Pylkkänen, 2017). Thus, considering this work as a whole, although increased LATL activation in compositional contexts may be accompanied by increases in other regions, the consistency of the LATL effect, particularly in MEG studies, and the developing functional characterization of the spatiotemporal pattern, makes this response an ideal springboard from which to assess the functional uniformity of different instances of composition.

1.2. *Prior work on the composition of words*

While the role of the LATL in composition between words has become increasingly clear, the possibility that this region also plays a similar role in composition within words has received relatively little attention. In the psycholinguistics literature, within-word composition has been proposed in accounts of complex word recognition, where it follows a stage of morphemic decomposition, or segmentation. Consistent with the obligatory decomposition model of Taft (2004), many studies have yielded results supporting the notion that decomposition takes place during the recognition of not only truly morphologically complex words (e.g., *hunt-er*), but also those words that can be completely parsed into existing morphemes, without being truly polymorphemic (e.g., *corn-er*; see Rastle & Davis, 2008 for a review of evidence). Taft (2004) provided evidence to support a framework wherein complex words whose meaning can be predicted from the meanings of their constituents (i.e., semantically transparent cases; *starlet, tombstone*), involve a decomposition and re-combination process to facilitate recognition. In contrast, when the meaning of the full word cannot be predicted from the constituents' meanings (i.e., semantically opaque cases; *doughnut*), it is posited that the decomposed morphemes activate a lemma-level representation of the whole word that is stored in memory, without a recombination process.

Alternative accounts of complex word recognition include parallel dual-route models (e.g., Baayen, Dijkstra, & Shreauder, 1997; Schreuder & Baayen, 1995), which postulate that the recognition of a complex word is underpinned by a "race" between a parsing mechanism that decomposes the word to its constituent morphemes and then re-combines them, and a full-word route that looks up the lemma based on the surface form. Such accounts have been supported by evidence showing that recognition processes can be influenced by characteristics of the whole word rather than, or in addition to, characteristics of the constituent morphemes (e.g., Bertram, Laine, & Karvinen, 1999; Häikö, Bertram, & Hyönä, 2011; Sites, Federmeier, & Christianson, 2016; see also Mankin, Thompson, Branigan, & Simner, 2016 for evidence from synesthesia), with more recent variants of the dual-route models stipulating that the extent to which each route is relied upon is influenced by many interacting sources of information, including properties pertaining to the full forms and constituents, as well as contextual and semantic cues (Kuperman, Bertram, & Baayen, 2008; Kuperman, Shreuder, Bertram, & Baayen, 2009). Results from electroencephalography (EEG) studies focused on the mismatch negativity (MMN) response component have also been interpreted as supporting dual-route accounts of complex word recognition (e.g., MacGregor & Shtyrov, 2013; see also Leminen, Leminen, Kujala, & Shtyrov, 2013; Hanna & Pulvermüller, 2014 for results related to whole-word access). However, the generalizability of these conclusions is unclear, given the low number of carefully chosen stimuli in each experiment, as well as the repetitive nature of the presentation of the crucial stimuli.

Interestingly, focusing on the attested patterns of electrophysiological responses in the literature, a common finding does arise in comparisons of various types of compound nouns across paradigms and languages. For instance, following the early amplitudes associated with the MMN, MacGregor and Shtyrov (2013) also observed an N400-like response to compound noun stimuli, which behaved in a manner consistent with an index of semantic integration of the constituents. That is, the component showed larger amplitudes in response to transparent compounds relative to opaque compounds, low frequency compounds relative to high frequency compounds (the former posited to be more likely to rely on a parsing route than a whole word route), and pseudo-compounds (i.e., novel compounds, which cannot rely on access to a full-form lexeme) relative to low frequency compounds. Previous electrophysiological results have found similar dissociations between varieties of compounding outside of the MMN paradigm. For instance, Koester, Gunter, and Wagner (2007) reported a slow negative shift over centro-parietal electrodes that was larger in response to German transparent compounds than opaque compounds, and showed increasing amplitude with increases in the reported plausibility of the constituent (presumed to index an integration cost; see Koester, Holle, & Gunter, 2009 for related evidence). Bai et al. (2008) observed a similar pattern in response to Chinese compounds, demonstrating that the response appeared to track a semantic composition process in particular, rather than a more general combinatory operation. The timing of these results is congruent with the MEG study of Fiorentino, Naito-Billen, Bost, and Fund-Reznicek (2014), which focused on English compounds, and showed a dissociation between responses to visually presented, lexicalized and novel compounds beginning approximately 400 ms after word onset, with increased amplitudes in response to novel compounds, presumed to index a greater reliance on compositional operations. Brooks and Cid de Garcia (2015) also found distinct amplitude increases in response to English transparent compounds in a window spanning 250–470 ms, with increases relative to monomorphemic nouns identified in the left anterior temporal lobe, and an absence of such a dissociation between opaque compounds and monomorphemic nouns.

Thus, although this later response component has been said to reflect a secondary, attention-mediated stage of processing in the context of dual-route models (MacGregor & Shtyrov, 2013), the consistency of the pattern, and its interpretation as indexing a semantic composition process, motivates the question of whether the underlying mechanism is shared with phrasal composition. Indeed, this was in part the motivation for the study of Brooks and Cid de Garcia (2015), which used comparisons between semantically transparent/opaque compound nouns and monomorphemic nouns to show localization of this effect to the left anterior temporal lobe, suggestive of a mechanism shared with similarly localizing phrasal composition effects (e.g., Bemis & Pylkkänen, 2011). In the present work, we have gone one step further by testing for overlap in the same design, and assessing how these processes interact in the construction of phrases. Further, motivated by parallels in the literature on decomposition and recombination during the processing of derived complex words (e.g., Fruchter & Marantz, 2015; Solomyak & Marantz, 2010), we extend this question to the domain of morphological derivation, and look for overlap with this process as well.

1.3. *Cumulative composition*

Finally, morphological composition, and the processing of morphological complexity more generally, does not typically take place in isolation. Rather, once an individual has processed a morphologically complex word in a phrase or sentence, it must be combined with the words that precede and follow it. To date, there has yet to be a comprehensive investigation addressing the question of whether, or how, the presence of word-internal composition or complexity influences the way in which a word enters into phrasal composition. For instance, do the specific processes required during the recognition of a semantically transparent compound such as *tombstone* cause it to combine with a preceding modifier (e.g., *granite tombstone*) in a way that is distinct from that which occurs in a phrase involving a monomorphemic noun (e.g., *granite fountain*)?

Consider for the moment, the typical MEG-localizing LATL response; if this represents a general phrasal composition mechanism that operates irrespective of the morphological complexity of the modified noun, then one would expect various types of morphologically complex nouns to elicit the same pattern of increased activity when they enter into phrasal composition. However, in light of the previous evidence that the LATL response indexes a process operating on semantic or conceptual representations, one would presume that it requires sufficient processing of the nouns themselves before phrasal composition can take place. In the case where a morphologically complex noun is proposed to require a combination of its constituent morphemes, this hypothesis would entail that the word-internal composition must occur before the phrase-level composition response appears. In other words, you must first put the pieces of the word itself together, before you can proceed to combine it with a preceding modifier. As a result, one would expect a difference in the timing of the LATL response during phrasal composition when modified nouns involve a within-word composition process, with LATL increases appearing at a delayed latency relative to phrasal composition involving monomorphemic nouns, which do not require the within-word operation. In particular, while the modification of a monomorphemic noun may elicit an effect 150–250 ms after noun onset, the same effect may appear with a delayed latency after the onset of a modified transparent compound or suffixed noun, due to the within-word process that is required. Alternatively, it could be that the presence of a word-internal composition process, or any type of morphological complexity, triggers a completely distinct processing stream, therefore eliminating the typical LATL response altogether.

1.4. *Current study*

The two experiments presented here were initially designed independently of each other, but are presented together, as they both apply the same basic formula to address the questions outlined above. In each experiment, we employed the minimal modifier-noun paradigm used in previous MEG studies (e.g., Bemis & Pylkkänen, 2011), while designing the stimuli to include specific types of morphological complexity that were intended to differ in the presence of within-word composition. We positioned our investigation of composition at a high level of generality, without controlling for factors such as the conceptual relationships present between composing elements, and the word-class of the morphemic constituents (e.g., adjective-noun compounding, noun–noun compounding, etc.). This was done with the hope that results across the heterogeneity will motivate future questions focused on finer differences in the phenomena.

Based on the model of Taft (2004) and related evidence (described above), we started with the assumption that instances of within-word composition are present in semantically transparent cases of morphological complexity, such as suffixed nouns (*star* + *let*) and transparent compound nouns (*tomb* + *stone*), where the meaning of the full word can be predicted from a combination of the constituents. In order to isolate neural responses associated with this composition process, we compared these complex noun conditions to monomorphemic nouns, containing the orthographic form of only one morpheme. Additionally, because this contrast could yield effects related to visual or orthographic demands that are increased during the recognition of bimorphemic words, we sought to include a second control condition that contained similar morpho-orthographic complexity (i.e., two morphemes present in the orthography of the words), but did not involve a within-word combination. For suffixed nouns, this was done with the inclusion of pseudo-suffixed nouns (e.g., *goblet*), whose surface strings can be separated into a stem that can function as a word on its own (*gob*), and a string that otherwise functions as a suffix (*-let*). For transparent compounds, a similar control condition was created using opaque compound nouns (e.g., *doughnut*), which contain two free morpheme constituents, but whose meanings cannot be predicted from a combination of these constituents ($dough + nut \neq doughnut$). Notably, while the morpho-orthographic constituents of pseudo-suffixed nouns have no semantic relationship with the full-form, this is not necessarily the case in opaque compound nouns (i.e., doughnuts are made from dough). Consequently, it is possible that some type of composition process may take place during the recognition of opaque compounds. With this in mind, the contrast of transparent and opaque compounds was intended to isolate responses specifically related to the conceptual combination that can yield the full meaning of semantically transparent cases.

In the primary analyses of this investigation, we first used a group-level functional localizer to isolate those regions in the left temporal lobe that underlie the basic LATL phrasal composition response, in the modification of monomorphemic nouns. We then tested whether these same regions were engaged in cases of morphological composition within-words. Experiment 1 was designed to isolate responses related to within-word composition in semantically transparent compounds (*tombstone*) as well as suffixed nouns (*starlet*), while experiment 2 was designed to isolate this process in suffixed nouns in a more targeted manner, using the pseudo-suffixed controls. Finally, by presenting nouns of each type of morphological complexity in both modified and unmodified contexts, we were also able to investigate whether phrasal composition involving each type of complexity relied on the same neural mechanisms as did that involving monomorphemic nouns or, if not, to test if distinct correlates could be identified.

2. Materials and methods

2.1. *Participants*

All participants were native English speakers with normal hearing, normal or corrected-to-normal vision, and were right-handed as diagnosed by the Edinburgh Handedness Inventory (Oldfield, 1971). All participants provided informed consent. Experiment 1 was conducted at the Neuroscience of Language Lab of New York University Abu Dhabi. 29 volunteers participated. 7 participants were removed from the analysis following data collection: 1 due to excessive movement during the MEG recording, 1 due to drowsiness during the MEG recording, and 5 due to un-removable electromagnetic interference. Therefore, a total of 22 participants were included in the experiment 1 analyses (9 males, 13 females, mean age = 25.4 years, SD = 9.03 years).

Experiment 2 was conducted across research facilities of the Neuroscience of Language Lab at New York University and New York University Abu Dhabi. 27 volunteers participated: 14 at the Abu Dhabi facility, and 13 at the New York facility. 4 participants were removed from the analysis following data collection due to un-removable electromagnetic interference. Therefore, a total of 23 participants were included in the experiment 2 analyses (9 males, 14 females, mean age = 25.0 years, SD = 7.11 years).

2.2. *Design*

Experiment 1 presented participants with sets of monomorphemic nouns (e.g., *rabbit*), suffixed nouns (e.g., *starlet*), semantically transparent compounds (e.g., *tombstone*), and semantically opaque compounds (e.g., *doughnut*), with each appearing once following a modifier, and once following a consonant string. The complete stimulus set consisted of 46 unique nouns in each condition of morphological structure, presented with both a modifier and a consonant string, yielding a total of 368 trials. Similarly, experiment 2 presented participants with nouns that were monomorphemic, suffixed, or pseudo-suffixed (e.g., *goblet*), with each noun again appearing once following a modifier, and once following a consonant string. In experiment 2 there were 50 unique nouns, each appearing with a modifier and a consonant string, for a total of 300 trials.

The trial structure, shown in Fig. 1, was similar to those used in previous MEG studies of modifier-noun composition (e.g., Zhang & Pylkkänen, 2015). Each trial began with the presentation a single fixation cross, followed by either a modifier or a consonant string, and then a target noun. Each word, as well as the fixation cross, was presented for 300 ms, with an interstimulus interval of 300 ms. At the Abu Dhabi facility, stimuli were presented on a screen approximately 80 cm away from the participant, and subtended a vertical visual angle of approximately 1.07°. At the New York facility, stimuli were presented on a screen approximately 45 cm away from the participant, and subtended a vertical visual angle of approximately .64°. In experiment 1, every trial was followed by a two-alternative forced-choice task, in which two words appeared side-by-side on the screen, and participants were asked to choose the one that most related to the word or phrase that they just read. In experiment 2, this task occurred on only 40% of the trials in order to reduce the duration of the experiment. We note that while participants could perform well on the task without paying attention to the modifying noun (see the example in Fig. 1), we have little reason to believe that participants would deviate from instruction and adopt this heuristic, as the matching task is likely to be considered quite easy to begin with. Further, such a behavior would presumably only bias against our chances of replicating the previous LATL composition-related effects, which have been identified in paradigms where such a strategy could not explain their presence (see Bemis & Pylkkänen, 2011; Bemis & Pylkkänen, 2013a; Zhang & Pylkkänen, 2015). In each condition, the correct answer was presented on the right side of the screen in half of the trials, and the left side of the screen in the other half. Participants indicated their answer by pressing one of two buttons with the index finger or middle finger of their left hand.

2.3. *Stimuli*

The stimuli sets for experiments 1 and 2 are summarized in Table 1, and lexical statistics are listed in Tables 2 and 3. The

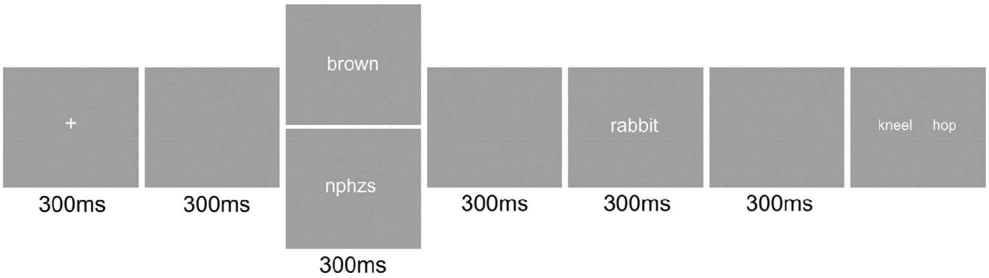

Fig. 1 – The structure of an individual trial. Participants were first presented with a fixation cross that was followed by either a modifier or an unpronounceable consonant string before the appearance of the target noun. Stimuli were on the screen for 300 ms, with an interstimulus-interval of 300 ms. Following the trial, participants were asked to choose which of two words most related to the phrase they had just read. In Experiment 1, this task occurred after every trial, while experiment 2 included the task after 40% of the trials.

Table 1 – Stimuli and Design. Monomorphemic, transparent compound, opaque compound, suffixed, and pseudo-suffixed nouns appeared in both modified and unmodified contexts across two experiments. The bottom row indicates the inclusion of a particular noun type in each experiment.

	Noun Type				
	Monomorphemic	Transparent Compound	Opaque Compound	Suffixed	Pseudo-suffixed
Modified	brown rabbit	granite tombstone	stale doughnut	blonde starlet	empty goblet
Unmodified	nphzs rabbit	xnlknlr tombstone	ccxkk doughnut	nlmnd starlet	lkmnh goblet
Experiment	1 & 2	1	1	1 & 2	2

complete sets of stimuli, as well as additional descriptions of the materials, can be found in the Appendix. We refer to all modifying words (the first word in each phrase) as modifiers. The second words, or phrasal heads, are simply referred to as nouns. For convenience, we refer to all nouns that are not monomorphemic as complex. The head nouns in all conditions were matched on measures of length, log-surface hyperspace analog to language (HAL; Lund & Burgess, 1996) frequency, and number of syllables with no statistically significant differences found in pairwise independent *t*-tests ($p > .10$) in each experiment. In each experiment, we sought to select modifiers that combined with these nouns in a natural way, while also balancing the conditions within each set on specific characteristics. It was not possible for both of these constraints to be perfectly satisfied, resulting in stimuli that were much more heterogeneous than those used in previous studies of minimal composition (cf. Bemis & Pylkkänen, 2011).

Monomorphemic, suffixed, and pseudo-suffixed nouns were selected from the English Lexicon Project (ELP; Balota et al., 2007). Monomorphemic nouns were accepted as stimuli if they were classified as containing only one morpheme in the ELP. Suffixed nouns were selected from those nouns that were classified as having 2 morphemes, and consisting of a free stem and a bound derivational suffix. Twenty-three unique suffixes were included in the experiment 1 materials, with each occurring twice. Eleven unique suffixes were included in the experiment 2 materials, with each occurring at least 3 but no more than 6 times in the set. Pseudo-suffixed nouns were selected by first extracting all monomorphemic nouns from the ELP, searching for those that ended with a pseudo-suffix, stripping the pseudo-suffix, and confirming that the remaining pseudo-stems were existing monomorphemic nouns on their own (see Appendix for additional filtering of pseudo-suffixed stimuli). We note that we did not seek to control the compatibility of the part of speech of the pseudo-stems with the pseudo-suffix, which could feasibly be thought to impact a mechanism attempting morphemic recombination.

Semantically transparent and opaque compounds were selected as a subset of the materials used in a previous study (Brooks & Cid de Garcia, 2015). In the previous work, a norming study split these items into the two conditions, based on the distinction that the meaning of the semantically transparent compounds could be derived from the straightforward combination of the two constituents, while the meaning of the opaque compounds could not. We note that semantic transparency is likely graded, rather than categorical, and that there may not be any absolutely transparent cases of morphological complexity. For our purposes, our dichotomy into transparent and opaque compound words was motivated by the desire to replicate the effects reported in previous literature. Compound conditions were not matched on the word class of their constituents (see Appendix for more details).

The sets of modifying words were chosen from the Corpus of Contemporary American English (COCA; Davies, 2008), and were matched for length, log-surface frequency, number of syllables and phrasal transition probability to the head noun, with pairwise independent *t*-tests revealing no significant differences across the levels of morphological complexity in each experiment ($p > .09$). Phrasal transition probability, based on data from COCA, was computed as the bigram frequency of

Table 2 – Experiment 1; lexical and phrasal statistics of the stimulus set. Means and standard deviations are reported for each lexical variable. Length, log surface Hyperspace-to-Analog frequency (Surface Frequency), and number of syllables were taken from the ELP. Transition probabilities were calculated based on values in COCA. Pairwise independent *t*-tests comparing Length, Surface Frequency, and Syllables for the modifiers and nouns separately, as well as Phrasal Transition Probability, showed no significant differences (all $p > .09$).

	Length	Log Surface Frequency	Syllables	Phrasal Bigram Frequency	Phrasal Transition Probability
Monomorphemic: Heads	8.35 (.640)	5.71 (1.25)	2.28 (.544)	4.93 (6.61)	.0013 (.0029)
Transparent: Heads	8.48 (.836)	5.82 (1.82)	2.15 (.420)	6.67 (17.1)	.0015 (.0039)
Opaque: Heads	8.52 (1.05)	5.71 (1.61)	2.15 (.363)	2.61 (3.80)	.0007 (.0021)
Suffixed: Heads	8.41 (1.45)	6.03 (2.22)	2.32 (.598)	4.91 (8.94)	.0007 (.0015)
Monomorphemic: Modifiers	5.11 (1.06)	9.53 (1.58)	1.41 (.580)		
Transparent: Modifiers	5.07 (1.29)	9.36 (1.91)	1.48 (.547)		
Opaque: Modifiers	5.13 (1.09)	9.51 (1.63)	1.48 (.547)		
Suffixed: Modifiers	5.48 (1.05)	9.34 (1.42)	1.57 (.583)		

Table 3 – Experiment 2; lexical statistics of the stimulus set. Means and standard deviations are reported for each lexical variable. Length, log surface Hyperspace-to-Analog frequency, and number of syllables were taken from the ELP. Transition probabilities were calculated based on values in COCA. Pairwise independent *t*-tests comparing each variable across the levels of morphological complexity, for the modifiers and nouns separately, found no significant differences (all $p > .10$). Base frequencies of the stems were comparable across the suffixed and pseudo-suffixed conditions, with no significant difference found in a *t*-test performed between the two. The average base frequency of suffixed nouns in Experiment 1 was 9.51 (SD = 1.60).

	Length	Log Surface Frequency	Syllables	Log Base Frequency	Bigram Frequency	Transition Probability
Monomorphemic: Heads	6.56 (.50)	7.76 (.51)	2.18 (.56)	n/a	27.08 (37.80)	.0040 (.0089)
Suffixed: Heads	6.60 (.99)	7.73 (2.02)	2.08 (.56)	9.01 (1.88)	26.06 (31.11)	.0041 (.0101)
Pseudo-suffixed: Heads	6.54 (1.15)	7.81 (1.96)	2.08 (.40)	8.87 (2.48)	24.78 (34.88)	.0037 (.0057)
Monomorphemic: Modifiers	4.92 (1.19)	9.84 (3.39)	1.38 (.57)			
Suffixed: Modifiers	5.04 (1.34)	9.99 (1.85)	1.46 (.58)			
Pseudo-suffixed: Modifiers	5.44 (1.54)	9.30 (1.93)	1.68 (.65)			

the modifier-noun phrase divided by the unigram frequency of the modifier when it appeared with the adjective part of speech tag, immediately preceding a noun. This was done in order to limit the frequency estimate of the modifiers to those instances where it could feasibly be followed by a noun in a phrase. For those modifiers that were not considered adjectives in COCA, we eliminated the part of speech tag criterion, using instead the count of times the word appeared immediately preceding a noun. Unigram frequency counts were estimated from the first 6000 unique collocates of the modifiers in order to make the look-up of their frequencies tractable. The phrases were not matched on bigram frequencies or the number of modifiers that were adjectives or nouns (see Appendix and Discussion for more details). ELP listed all of the selected modifiers as monomorphemic.

For each modifier, a length-matched, unpronounceable consonant string was generated for the corresponding unmodified condition. This contrast results in a difference in the number of words encountered on each trial in the modified and unmodified conditions. While this could potentially mean that composition-related responses could be confounded by effects related to the processing of two words in the modified condition, previous results that have used two word non-combinatory controls (e.g., a list of two words; Bemis & Pylkkänen, 2011) have demonstrated that the LATL's response to phrasal composition cannot be attributed to such a difference. In the present experiments, the introduction of list conditions would have added considerable length to the experimental procedures. As an alternative, we opted to omit this contrast, and instead rely upon the previous results to inform our interpretations of effects related to phrasal composition.

2.4. *Procedure*

Institutional Review Boards overseeing research involving human subjects at New York University and New York University Abu Dhabi approved all experimental procedures involved in this research. Before beginning the MEG recording, each participant's head shape was digitized with a Polhemus FastSCAN system (Polhemus, Vermont, USA). In addition to the shape of each participant's head, the positions of three fiducial landmarks (the nasion and left/right tragi) and five head position indicator (HPI) coils (located on the forehead, and at pre-auricular points) were recorded for later use in the alignment of coordinate systems. Following the head scan process, each participant completed a short practice session at a desktop computer. The practice instructions were identical to those given in the main experiment. The practice session consisted of ten trials involving five monomorphemic nouns in both modified and unmodified contexts. Participants received feedback ("Correct" or "Incorrect") after every practice trial. Neither the nouns nor the modifiers used in the practice sessions were repeated in the main experiment.

Participants completed the experiment while laying supine in a dimly lit magnetically shielded room (MSR). At the Abu Dhabi research facility, a 208-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan) recorded the continuous MEG data throughout the experiment. At the New York research facility, a 156-channel axial gradiometer whole-head system (Kanazawa Institute of Technology, Kanazawa, Japan) recorded the continuous MEG data. Data were collected at a sampling rate of 1000 Hz. During data acquisition, the data were filtered using high and low-pass filters of .1 Hz and 200 Hz, respectively. The positions of the five HPI coils, attached to the forehead and pre-auricular points, were measured relative to the MEG sensors before and after the recording. Participants typically required 20–30 min to complete the entire experiment.

2.5. *MEG data pre-processing*

The MEG data were first noise-reduced using the Continuously Adjusted Least-Squares Method (Adachi, Shimogawara, Higuchi, Haruta, & Ochiai, 2001) implemented in the MEG160 software (Yokogawa Electrical Corporation and Eagle Technology Corporation, Tokyo, Japan). All remaining data pre-processing steps, and subsequent analyses, were performed with MNE-python (v. .12; Gramfort et al., 2013, 2014) and Eelbrain (v.0.24.6). Due to extreme low frequency electromagnetic noise in the New York City environment, it was not possible to analyze the MEG data collected at the New York facility without applying a high-pass filter at 1 Hz. In order to reduce the asymmetry in the dataset, the same high-pass filter was applied to all experiment 2 data collected at the Abu Dhabi facility. Notably, the application of a 1 Hz high-pass filter has been proposed to lead to spurious effects in evoked responses (e.g., Tanner, Morgan-Short, & Luck, 2015). In this light, the experiment 1 data would be considered less likely to contain artificial effects. For this reason, when considering effects present in the Experiment 2 data, we were careful to only interpret as meaningful those that were replicated in the Experiment 1 data, which can therefore be considered stable across filter settings. Data collected in both experiments and at both facilities were low-pass filtered at 40 Hz. Following filtering, independent component analysis was applied to the continuous data for each participant. The components were visually inspected, and those fitting the profile of known artifacts (e.g., eye blinks, movement-related activity, and well-characterized external noise sources), were removed. The MEG data were then segmented into epochs spanning 100 ms before the onset of the initial word to 1200 ms following, and baseline corrected relative to the initial 100 ms pre-stimulus interval (note that 0 ms in the following analysis and figures corresponds to the onset of the target nouns). Epochs containing amplitudes greater than an absolute threshold of 2000 fT were removed before further preprocessing. This resulted in the rejection of 1.45% of all epochs. Finally, for each participant, all remaining epochs were averaged in each condition to generate an evoked response at the sensor level.

Cortically constrained minimum-norm current estimates were computed for each participant's evoked responses. The “fsaverage” brain, available with FreeSurfer (https://surfer.nmr.mgh.harvard.edu/), was scaled to match each participant's digitized head shape, and then fit to that head shape by aligning the positions of the nasion and pre-auricular points with those recorded in the digitization. A source space of 2562 source points per hemisphere was generated on the cortical surface for each participant. The boundary element model was employed to compute a forward solution, explaining the contribution of activity at each source to the magnetic flux at the sensors. Channel-noise covariance was estimated based on 100 ms intervals prior to each artifact-free trial. Given the forward solution and channel noise-covariance estimates, the inverse solution was computed with “free” orientation on the MRI coordinate system, and then applied to each evoked response. The source estimates were unsigned, meaning that the polarity (i.e., outgoing and incoming with respect to the cortical surface) was ignored, and calculated using noise-normalization to yield dynamic statistical parameter maps (dSPM; Dale et al., 2000).

2.6. *Statistical analysis*

2.6.1. *Functional localizer*

We first performed a group-level functional localizer analysis to isolate those portions of the left temporal lobe that showed increased activity in response to phrasal composition. This analysis, as well as all subsequent analyses except the spatial overlap tests, was performed on condition averages from individual subjects, which are referred to as evoked responses. The functional localizer consisted of a spatiotemporal, cluster-based permutation test (see Maris & Oostenveld, 2007) comparing evoked responses to modified and unmodified monomorphemic nouns. The general procedure for the cluster-based test was as follows: A univariate test statistic was calculated at each time point, at each discrete source on the cortical surface, within the analyzed region and time window. Clusters were then formed from test statistics that were contiguously significant through time and space, at a level of $p < .01$. We selected this rather strict threshold for all spatiotemporal tests in order to identify focal regions in the left temporal lobe that correspond to the sites of peak sensitivity. Only those clusters that contained a minimum of ten sources and spanned at least 10 ms were considered any further. For each observed cluster that met these criteria, we calculated the sum of all test statistics contained within it to obtain a cluster-level statistic. To determine the reliability of these clusters relative to variance in the data, we employed a randomization procedure to approximate the null distribution (Maris & Oostenveld, 2007). This involved randomly shuffling the assignment of condition labels to each participant's evoked responses, repeating the mass univariate test and clustering procedure within the analysis window, and forming a distribution from the largest cluster-level statistic that was observed in each re-shuffle. This procedure was performed 10,000 times, resulting in a distribution of 10,000 values. Each of the truly observed clusters was subsequently assigned a p-value based on the proportion of the surrogate null distribution greater than its own cluster-level statistic. The center of mass for each spatiotemporal cluster in Montreal Neurological Institute (MNI) coordinates was calculated by weighting the position of each constituent source by its mean t-value over the time window of the cluster.

Based on the previous results identifying the LATL composition response (e.g., Bemis & Pylkkänen, 2011; Westerlund & Pylkkänen, 2014), we constrained our functional localizer analysis window to 150–350 ms post-noun onset. The left temporal lobe was defined based on the PALS-B12 parcellation (Van Essen, 2005). The test calculated at each source-time point was a paired t-test, evaluated as one-tailed for an increase in the modified condition. In order to maximize statistical power and identify those regions of the left temporal lobe that were most consistently recruited during basic phrasal composition, we combined both experiment 1 and experiment 2 subjects in this initial analysis. Post-hoc inspection revealed similar patterns in each dataset independently (see Results section). All subsequent analyses were performed on each experimental dataset separately.

2.6.2. *fROI analyses of morphological complexity*

After identifying LATL regions that were modulated by phrasal composition, we then adopted these as functional regions of interest (fROIs) and tested whether they were similarly sensitive to morphological composition using temporal, cluster-based comparisons of each noun condition. These temporal tests were procedurally similar to the spatiotemporal test described above, with the exceptions that each participant's evoked responses were time-courses averaged over the sources in the fROI, therefore eliminating the spatial element to the test, and that we adopted the more conventional cluster-forming threshold of $p < .05$, as we were no longer emphasizing the identification of focal spatial clusters. We first used pairwise comparisons between the related complex cases themselves (transparent *vs* opaque compounds and suffixed *vs* pseudo-suffixed), and each complex case and the monomorphemic nouns. On the basis of the model of Taft (2004), generalizing to derivation and compounding, and previous MEG studies focused on morphemic recombination (i.e., Brooks & Cid de Garcia, 2015; Fruchter & Marantz, 2015), we expected correlates of a morphological composition process to manifest as increases in the suffixed and transparent compound noun conditions, relative to the pseudo-suffixed and opaque compound nouns, respectively, and the monomorphemic noun conditions.

In addition to these pairwise tests, we also performed a cluster-based intersection test for source/time patterns that showed increased amplitudes in response to transparent cases of complexity (transparent compounds and suffixed nouns) relative to both monomorphemic nouns and their complex counterparts (opaque compounds and pseudo-suffixed nouns, respectively). In this manner, we intended to more directly account for increases due to the presence of morpho-orthographic complexity alone, which should be shared between transparent/opaque compounds, and suffixed/pseudo-suffixed nouns. This involved performing the two relevant *t*-test comparisons (in experiment 1: transparent *vs* monomorphemic, transparent *vs* opaque; in experiment 2: suffixed *vs* monomorphemic, suffixed *vs* pseudo-suffixed) at every point in the time course/spatiotemporal window, identifying the lesser of the two t-statistics that resulted from these tests, forming clusters from said statistics that met the significance threshold ($p < .05$ in temporal tests, $p < .01$ in spatiotemporal tests), and then testing for the statistical significance of these clusters. These tests thus provided a more targeted method of isolating clusters of activity that matched the assumed pattern of within-word composition, without the need for inferences across multiple individual tests of statistical significance. Once significant clusters were identified in this intersection analysis, we performed the remaining contrast of the other complex condition and monomorphemic nouns (opaque compound *vs* monomorphemic, pseudo-suffixed *vs* monomorphemic) on mean activity within the cluster, since a correlate of a within-word composition process unique to the transparent cases should not show sensitivity to the difference between these two conditions.

The analyses of suffixation were performed separately for each experiment. This was chosen over the alternative of combining the data from the two studies, as was done in the functional localizer analysis, due to the considerable differences between the suffixed noun stimuli in each experiment (see Appendix). While similar differences were present between the sets of monomorphemic nouns in each experiment, the consistency of the previously identified LATL response to phrasal composition motivated our decision to combine the data in the functional localizer analysis.

All tests for morphological composition were evaluated as one-tailed for greater activation in response to the complex nouns relative to monomorphemic nouns, and transparent compounds/suffixed nouns relative to their non-monomorphemic counterparts. All analyses were performed on activity within a 150–350 ms window, and 350–600 ms window. For every test, the significance of each observed cluster was determined by comparison of that cluster's statistic to the null distribution generated via randomization in the corresponding region and time window.

2.6.3. *Spatial divergence between phrasal and morphological composition*

Where potential correlates of phrasal and morphological composition were identified in regions outside of the fROIs, we tested for the presence of reliable spatial divergence on an individual subject basis. This was done by applying each subject's inverse solution to their individual trial data in order to generate source estimates for responses to each stimulus in the experiment. Within the data from each subject, we then calculated t-statistics at each source-time point in the left temporal lobe, corresponding to the contrast between modified and unmodified monomorphemic conditions in the time window of the effects identified in the functional localizer. Likewise, we also calculated t-statistics at each source-time point in the left temporal lobe for the contrast revealing the morphological composition effect, within the window identified in the group-level analysis. After averaging the statistics over their corresponding time windows, we identified the highest ten percent of mean statistics from each contrast (49 out of 499 possible sources in the left temporal lobe), and counted the number of overlapping sources amongst the two contrasts (i.e., those sources that show t-values in the highest ten percent in both comparisons). While the choice of a ten percent cut-off was arbitrary, follow-up analyses showed that the conclusions were robust across multiple choices of cut-offs ranging from five to twenty-five percent. A magnitude for group-level overlap was calculated by averaging the number of overlapping sources across subjects. If the two processes truly relied upon spatially distinct loci, we would expect a low degree of spatial overlap across the group. A permutation procedure with 1000 iterations was used to construct a surrogate null distribution for overlap by randomly re-arranging the condition labels within each subject's data, and re-calculating the group-level overlap score from the permuted data. The observed overlap was assigned a *p*-value based on the proportion of the resulting distribution of overlap scores that was less than the observed value. If the *p*-value was less than .05, we considered the divergence to be statistically significant.

2.6.4. *Phrasal composition involving complex nouns*

Finally, in order to assess whether phrasal composition involving each type of complex noun recruited the same temporal lobe regions as phrasal composition involving monomorphemic nouns, we again tested activity localized to the fROIs for increases related to composition. For each type of complex noun, we used a temporal, cluster-based permutation test to compare responses in the modified and unmodified conditions. The univariate test in each case was a paired *t*-test, which was evaluated as one-tailed for increases related to modification. Tests were conducted in the original 150–350 ms time window, as well as the later 350–600 ms window. If no significant temporal clusters were found, we expanded the analysis to the entire left temporal lobe and assessed whether reliable spatiotemporal clusters formed in distinct sub-regions. In the event that no reliable clusters were found, we presented the full time-course of insignificant spatiotemporal clusters for visual inspection. In the case of suffixed nouns, we performed both the combined and separate analyses.

3. Results

3.1. *Behavioral results*

The mean participant accuracy in each condition is reported in Tables 4 and 5, for experiments 1 and 2, respectively. Accuracy was consistently high, demonstrating successful completion of the task. The average participant accuracy was 96% (SD = 2.91%) in experiment 1, and 86.5% (SD = 6.2%) in experiment 2. We attributed the slightly lower accuracy values in experiment 2 to the reduced frequency of the task, which was present on only 40% of the trials. Since the purpose of the task was solely to ensure that participants attend to the stimuli, the behavioral data were not analyzed further.

3.2. *MEG results*

Data from both experiments are available at the Open Science Framework (http://doi.org/10.17605/OSF.IO/7BD29).

3.2.1. Functional localizer: phrasal composition involving monomorphemic nouns

As shown in Fig. 2, the functional localizer revealed a significant cluster that covered a major portion of the left anterior temporal lobe, stretching from the inferior temporal gyrus to the lateral sulcus. The cluster's duration spanned 162–299 ms post noun-onset ($p = .0002$). Inspection of the t-values within the spatial boundaries of the cluster revealed two locations with peak sensitivity to the contrast of modified and unmodified conditions. One of these was in the superior temporal gyrus and lateral sulcus, while the other was located in the inferior temporal gyrus. Since these two peaks of sensitivity were connected by only a narrow band of sources, the decision was made to separate the two foci. To do so, we dropped the cluster-forming p-value to half of its original value ($p = .005$), therefore only including those sources that met a higher t-value threshold. A repetition of the spatiotemporal test confirmed that the harsher threshold split the larger cluster into two. One of these two clusters appeared 170–288 ms [$p = .0006$, MNI: (−40, −26, 4)] after noun onset, and captured anterior sections of the superior temporal gyrus (aSTG) and the lateral sulcus. The second appeared 202–298 ms [$p = .0054$, MNI: (−53, −25, −31)] after noun onset, and was primarily constrained to the anterior section of the inferior temporal gyrus (aITG). When the functional localizer analysis was performed on the later time window (350–600 ms), no significant clusters were found ($p > .30$). Inspection of the results in each experiment separately demonstrated a replication of this pattern across the two datasets, with both showing statistically significant or nearly significant spatiotemporal clusters in the anterior sections of the superior [Experiment 1: 168–262 ms, $p = .0181$, MNI: (−35, −24, −1); Experiment 2: 220–282 ms, $p = .0546$, MNI: (−47, −22, 7)] and inferior temporal gyri [Experiment 1: 168–257 ms, $p = .0740$, MNI: (−52, −32, −28); Experiment 2: 248–309 ms, $p = .0581$, MNI: (−48, −15, −37)]. The appearance of spatially distinct clusters in each experiment, using the original p-value threshold, reinforced our decision to split the larger cluster into two distinct regions. The spatial extents of the two clusters from the combined functional localizer analysis were adopted as fROIs for all of the following analyses.

Table 4 – Experiment 1; mean participant accuracy in each condition.

		Phrasal Modification	
		Modified	Unmodified
Morphological Complexity	Monomorphemic	.952	.942
	Transparent Compound	.981	.981
	Opaque Compound	.966	.939
	Suffixed	.973	.977

Table 5 – Experiment 2; mean participant accuracy in each condition.

		Phrasal Modification	
		Modified	Unmodified
Morphological Complexity	Monomorphemic	.883	.893
	Suffixed	.898	.774
	Pseudo-suffixed	.872	.870

3.2.2. Morphological complexity

In the experiment 1 analyses, pairwise comparisons between the compound and monomorphemic conditions revealed one significant cluster in the fROIs associated with phrasal composition. As shown in Fig. 3, this was found in the aSTG fROI, in the contrast of opaque compounds and monomorphemic nouns, capturing increased amplitudes in response to the opaque compounds, 195–245 ms ($p = .0082$) following word onset. No significant clusters were found in the contrasts of transparent *vs* opaque compounds ($p > .20$) or transparent compounds *vs* monomorphemic nouns ($p > .20$). Spatiotemporal versions of these tests performed across the entire temporal lobe also failed to find any significant clusters ($p > .10$), with the exception of the contrast of opaque compounds and monomorphemic nouns, which identified a significant spatial cluster, matching the timing of the fROI result [190–255 ms, $p = .0347$, MNI: (−50, −9, −17)], though the spatial peak of the effect was positioned in the middle temporal gyrus.

The more targeted intersection analysis in the fROIs failed to identify any significant clusters ($p > .10$). We then performed the intersection analysis of compounding in a spatiotemporal test of activity across the entire left temporal lobe. This revealed a single cluster located in the temporal pole, as shown in Fig. 3, which captured increases in response to transparent compounds relative to both opaque compounds and monomorphemic nouns. The cluster spanned 457–475 ms, with a p-value sitting just above the threshold for

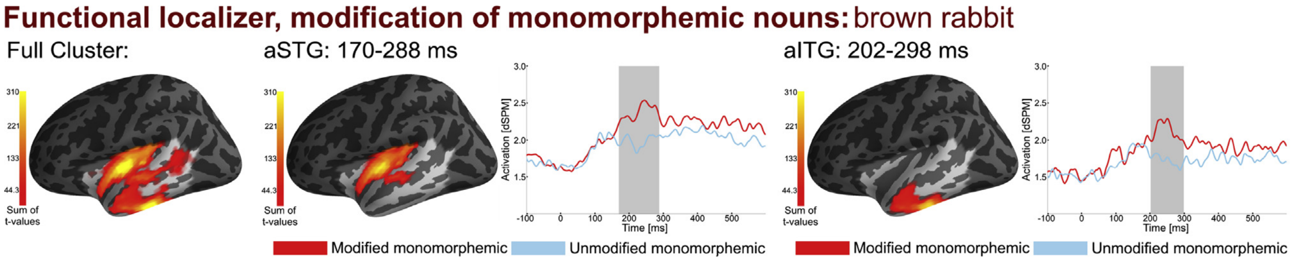

Fig. 2 – Regions of the left temporal lobe that showed increased activation in response to modified monomorphemic nouns, relative to unmodified monomorphemic nouns, were first identified in a functional localizer analysis. Data from both experiments were included in a single analysis in order to maximize power and capture the spatiotemporal patterns that were consistent across both. The analysis identified a large cluster stretching across the left anterior temporal lobe. To separate the two foci of the cluster, we increased the statistical threshold for inclusion, and repeated the analysis. This yielded one cluster in the anterior superior temporal gyrus and lateral sulcus (aSTG; 170–288 ms) and another in the anterior inferior temporal gyrus (aITG; 202–298 ms). The spatial extents of these two clusters were subsequently adopted as functional regions of interest in the following analyses.

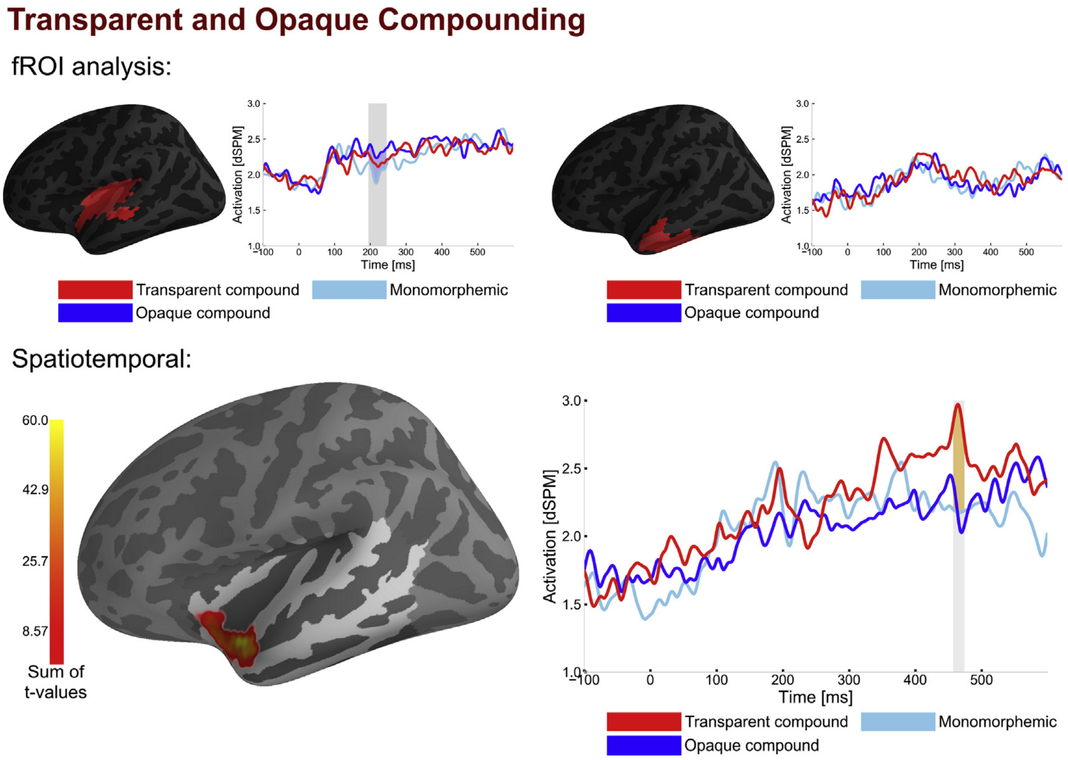

Fig. 3 – Transparent and opaque compounding. Top: pairwise temporal cluster-based tests were performed between the time-courses of activity in the fROIs. This revealed a single significant cluster in the contrast of opaque compounds and monomorphemic nouns, spanning 195–245 ms ($p = .0082$). Bottom: the expansion of the intersection analysis to a spatiotemporal test of activity across the left temporal lobe revealed a nearly significant cluster in the left temporal pole, spanning 457–475 ms ($p = .0502$). This cluster captured increased amplitudes in response to transparent compounds relative to both opaque compounds and monomorphemic nouns.

significance [$p = .0502$, MNI: (−45, 4, −22)]. The follow-up t-test between mean amplitudes in response to opaque compounds and monomorphemic nouns, within the cluster, revealed an absence of a significant difference between the two [$t(21) = -.12$, $p = .546$].

We subsequently performed a test of the spatial dissociation between the transparent compound effect identified in the intersection analysis and the phrasal composition effect identified in the functional localizer. For each subject, we first identified those sources that had the highest ten percent (49 sources) of t-statistics in the contrasts of modified and unmodified monomorphemic nouns, and unmodified transparent compounds and opaque compounds, in the time windows of the group-level effects. We then assessed the degree of overlap amongst these sources from each contrast, and averaged across subjects to generate a group-level measure of overlap. The group-level overlap (3.59 out of 49 sources) was not significantly less than what would be expected based on the permutation ($p = .150$), failing to provide evidence for reliable divergence. The same conclusion was reached when using the contrast of unmodified transparent compounds and monomorphemic nouns (i.e., the other contrast in the intersection analysis; group-overlap = 2.18 out of 49 sources, $p = .732$).

The fROI results pertaining to suffixation are shown in Fig. 4. The experiment 1 comparisons of responses to unmodified suffixed and monomorphemic nouns revealed a significant cluster in the aSTG fROI, spanning 225–257 ms post-word onset ($p = .0308$). No significant clusters were found in the aITG fROI ($p > .09$). In experiment 2, pairwise comparisons revealed a significant cluster in the aITG region, which captured increased responses to suffixed nouns relative to monomorphemic nouns, between 186 and 203 ms post-word onset ($p = .0489$). Notable for its similarity in location and timing, a near-significant cluster was also found in the aITG in the comparison of suffixed nouns to pseudo-suffixed nouns, between 188 and 203 ms ($p = .0792$). No other significant clusters were identified in the pairwise tests (all $p > .10$) in the fROIs. In spatiotemporal versions of these tests, only the comparison of suffixed and pseudo-suffixed conditions revealed a significant cluster. This cluster was found in the inferior temporal and fusiform gyri, approximately matching the location and timing of the fROI result [180–214 ms, $p = .0222$, MNI: (−40, −41, −23)]. The intersection analysis confirmed the pattern of results in the pairwise tests, with a single significant cluster identified in the aITG fROI, between 188 and 201 ms ($p = .0079$), which was also found in the spatiotemporal version of this test (182–201 ms, $p = .0088$, MNI: [−44, −33, −22]). Follow-up t-tests between responses to the pseudo-suffixed and monomorphemic nouns showed no significant differences in those clusters identified in the intersection-based analyses

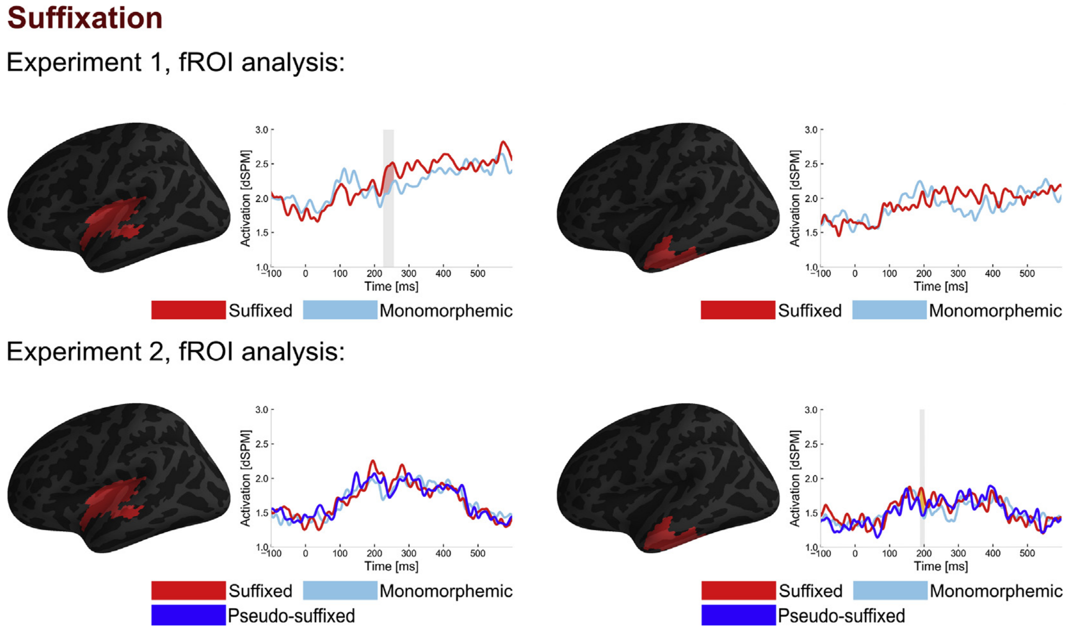

Fig. 4 – Suffixation. Cluster-based permutation tests of activity localized to each fROI revealed significant modulation of response amplitude by the type of the noun encountered, though the localization of this effect was not consistent across the two experiments. The results of each experiment are presented separately. In experiment 1, clusters were formed from point–wise statistics corresponding to increased amplitudes in response to suffixed nouns relative to monomorphemic nouns, revealing a significant cluster in the aSTG between 225 and 257 ms (top-left; $p = .0308$). In experiment 2, pairwise cluster-based tests between the conditions revealed overlapping effects in the aITG fROI (bottom-right). A significant cluster was found in the comparison of suffixed and monomorphemic nouns (186–203 ms, $p = .0489$), and a nearly significant cluster was found in the comparison of suffixed and pseudo-suffixed nouns (188–203 ms, $p = .0792$). The intersection-based test confirmed these results, revealing a significant cluster in the aITG between 188 and 201 ms ($p = .0079$), shaded in the bottom–right plot. No significant clusters were found in the remaining fROI in each experiment.

[temporal test: t(22) = .04, $p > .95$; spatiotemporal test: t(22) = .39, $p > .70$].

3.2.3. *Phrasal composition involving complex nouns*

As can be seen in Fig. 5A, the fROI analysis of phrasal composition involving semantically transparent compounds (e.g., *granite tombstone*), revealed significant clusters of increased activation in both the aSTG and aITG. In the aSTG, the first significant cluster was found between 212 and 258 ms ($p = .0318$) post-noun onset, and the second was found between 528 and 600 ms ($p = .0060$). In the aITG, a sole significant cluster was found between 521 and 600 ms ($p = .0037$). As shown in Fig. 5B, the fROI tests did not reveal any clusters in the comparisons of responses to modified and unmodified suffixed nouns, in either experiment. Similarly, the expansion of these tests to spatiotemporal analyses did not yield any significant clusters in either dataset ($p > .50$). The full time course of clusters that appeared in the temporal lobe, but did not reach the threshold for statistical significance, can be seen in Fig. 5B. We also note that when data were combined from the two experiments, the modification of suffixed nouns still failed to elicit any significant clusters in either fROI, or in spatiotemporal tests ($p > .30$).

Finally, instances of phrasal composition involving an opaque compound, or a pseudo-suffixed noun, did not elicit any statistically significant clusters in the fROIs (all cluster *p*-values > .06). In both cases, however, near-significant clusters were observed in at least one fROI. On the modification of opaque compounds, a trending cluster was found in the aITG between 294 and 322 ms ($p = .0639$). On the modification on pseudo-suffixed nouns, a trending cluster was found in the aSTG between 244 and 272 ms ($p = .0752$). Follow-up expansions of these tests to the entire temporal lobe failed to reveal any significant spatiotemporal clusters of increased activity related to composition (all cluster *p*-values > .50). The complete time courses of evoked responses in the fROIs, and insignificant clusters in the temporal lobe, are displayed in Fig. 6.

3.2.4. *Follow-up analyses*

As an alternative to using fROIs from the combined functional localizer analysis one could define experiment-specific fROIs using the contrast between modified and unmodified monomorphemic nouns from each dataset individually. In a follow-up to our original tests, we did just this, using the significant and nearly significant aSTG and aITG clusters identified within each experiment to check for similarly localizing effects related to compounding (experiment 1) and suffixation (experiments 1 and 2), as well as the modification of each type of complex noun.

The results of these analyses, which were performed in the same manner as previously described, showed the same pattern of results in both experiments. In experiment 1, pairwise comparisons of compounds and monomorphemic nouns found only a single significant cluster in the aSTG capturing increases in response to opaque compounds relative to monomorphemic nouns (195–229 ms, $p = .0200$; all other *p*-values > .08), comparisons of suffixed and monomorphemic nouns found a single significant cluster in the aSTG (227–257 ms, $p = .0317$; all other *p*-values > .25), and the modification of each type of noun only elicited significant effects in transparent compounds, in the same manner as the original analysis (aSTG: 212–252 ms, $p = .0348$, 530–600 ms $p = .0055$; aITG: three near-contiguous clusters spanning 394–600 ms, all $p < .005$; all other *p*-values > .07). In experiment 2, the clusters previously identified in the pairwise comparisons no longer reached the threshold for statistical significance, but the same patterns were present in the data (suffixed > monomorphemic in aITG: 190–203 ms, $p = .1550$; suffixed > pseudo in aITG: 190–201 ms, $p = .1978$). The result identified in the intersection analysis, capturing increased amplitudes unique to suffixed nouns, remained significant (aITG: 190–201 ms, $p = .0370$). No new significant effects were identified in the experiment-specific fROI analyses in either experiment (all remaining *p*-values > .05).

Lastly, in a post-hoc analysis motivated by dual-route models of complex word recognition, we attempted to parse apart a reliance on full-form and decomposition/re-combination routes. To do so, we split our compound stimuli in experiment 1 into equal groups of high and low frequency, transparent and opaque compounds (mean log frequencies: high transparent = 7.27, low transparent = 4.36, high opaque = 6.97, low opaque = 4.45), following the reasoning that frequent items are more likely to benefit from, and thus promote the existence of, full-form lexemes that can be accessed without decomposition and re-combination (as done by MacGregor & Shtyrov, 2013). The monomorphemic noun condition of experiment 1 was also split on frequency (mean log frequencies: high monomorphemic = 8.17, low monomorphemic = 4.73). Pairwise comparisons of the noun types within each frequency bin failed to find any significant temporal clusters in either fROI (all cluster *p*-values > .20), and the same was true of spatiotemporal versions of these tests performed across the temporal lobe ($p > .10$). The intersection-based tests in each frequency bin revealed an absence of any clusters meeting the minimum space and time criteria in either fROI, and the same was true in spatiotemporal tests across the temporal lobe.

4. Discussion

The first goal of this investigation was to assess whether morphologically complex words exhibit similar neural correlates of composition as phrases. To achieve this, we first compared evoked responses to modified and unmodified monomorphemic nouns, finding a widespread increase across the left temporal lobe in response to phrasal composition. Within this cluster, two distinct spatial foci were found in inferior and superior sections the lobe, both of which showed increased activity in response to modified monomorphemic nouns between 200 and 300 ms post-onset. The appearance of two foci was consistent across each of the independent experiments presented here, in the face of considerable differences in characteristics such as lexical frequency of the head nouns, bigram frequency of the phrases, transition probability from modifier to noun, and the number of modifiers that were adjectives versus nouns (see Tables 2 and 3).

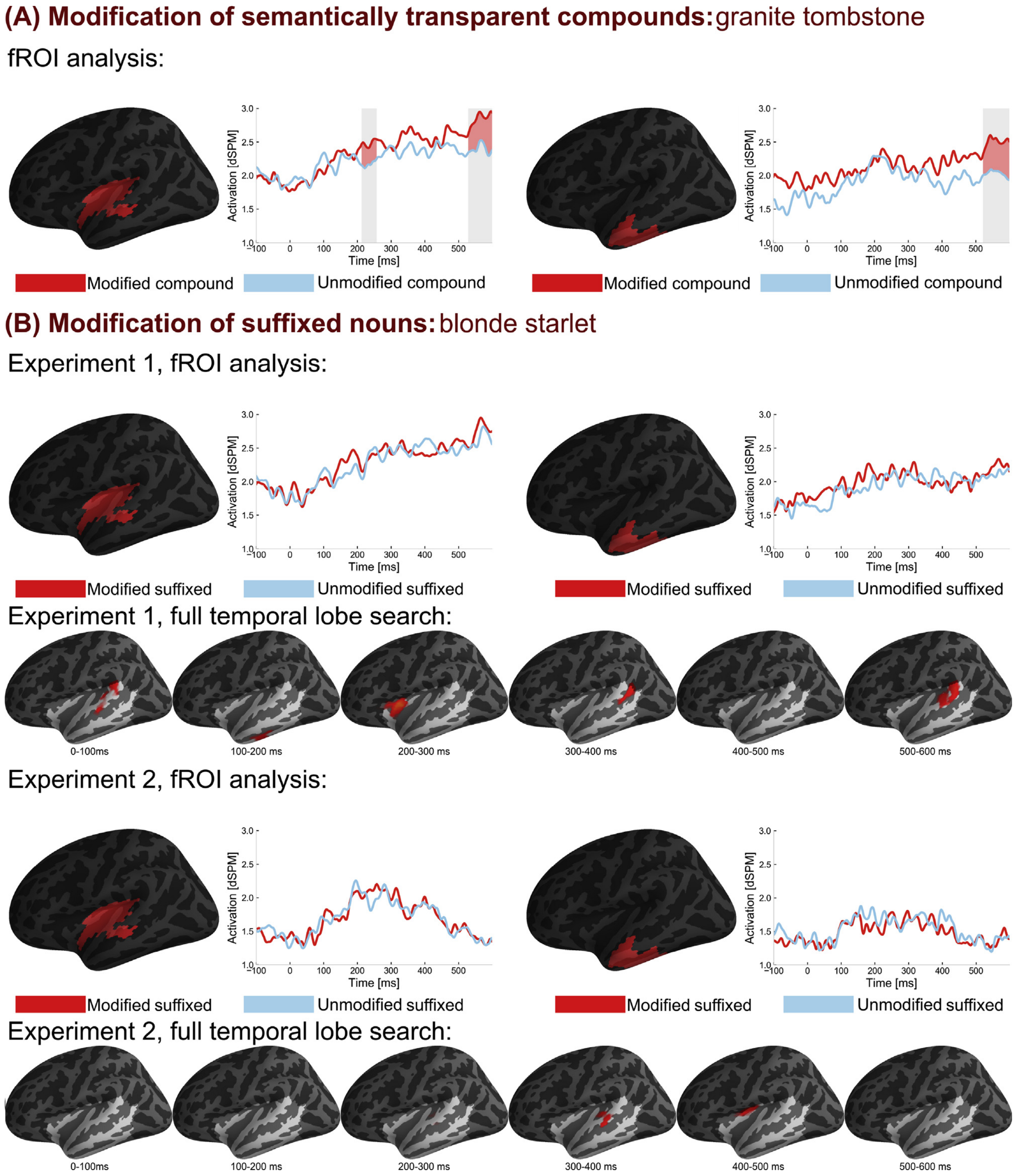

Fig. 5 – **(A) The modification of semantically transparent compounds. Cluster-based permutation tests performed on activity in response to modified and unmodified transparent compounds revealed significant clusters in both of the fROIs. In the aSTG, the first cluster appeared between 212 and 258 ms, ($p = .0318$), and the second between 528 and 600 ms ($p = .0060$). In the aITG, a sole significant cluster was found between 521 and 600 ms ($p = .0037$). (B) The modification of suffixed nouns. Across both experiments, no clusters were found in either of the two fROIs. The expansion of the test to the entire left temporal lobe failed to reveal any significant clusters.**

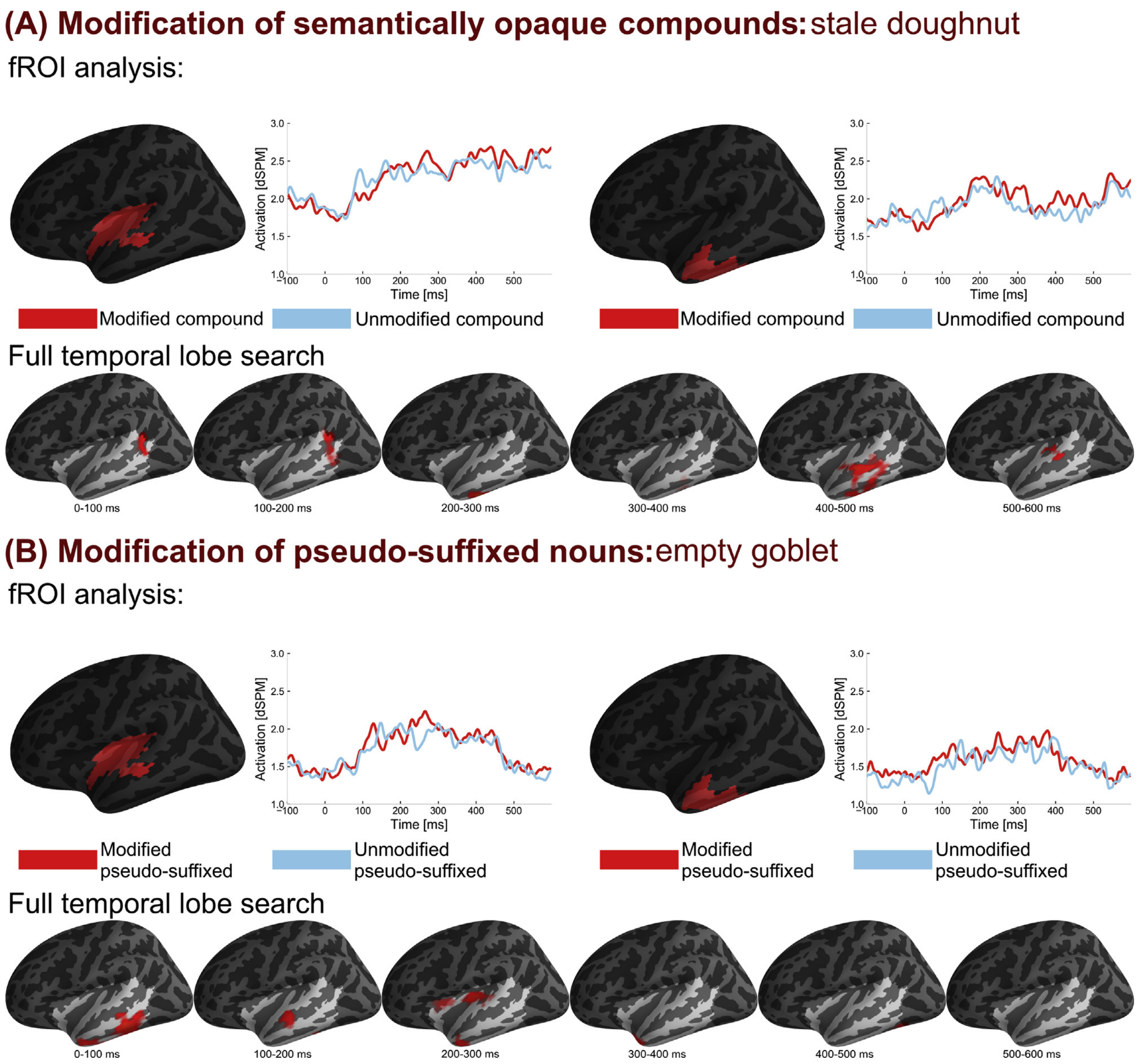

Fig. 6 – **(A) The modification of semantically opaque compounds. Cluster-based permutation tests did not reveal any significant clusters in either fROI, nor across the left temporal lobe. (B) The modification of pseudo-suffixed nouns. Cluster-based permutation tests did not reveal any significant clusters in either fROI, nor across the left temporal lobe.**

Having characterized these two regions as being sensitive to phrasal composition, we then tested the sensitivity of each to the presence of various types of morphological complexity, with specific interest in cases of complexity that have been proposed to involve a morphemic re-combination process. Transparent compounding (e.g., *tombstone*) did not modulate activity in these regions but, instead, was associated with increased activation in a neighboring section of the temporal pole, between 450 and 475 ms post word-onset. Suffixation (e.g., *starlet*), on the other hand, did modulate activity in the regions associated with phrasal composition, but not consistently across the two experiments, thus limiting our ability to draw confident conclusions. This modulation was found first in the results of experiment 1, where the more superior aSTG region associated with phrasal composition showed increased response amplitudes to suffixed nouns relative to monomorphemic nouns. In experiment 2, this result was not replicated, however an increase in response to suffixed nouns relative to both pseudo-suffixed and monomorphemic nouns was identified in the aITG region associated with phrasal composition. Pairwise comparisons between each type of complexity and the monomorphemic condition in each experiment failed to identify any consistent increases that could be related to the processing of morpho-orthographic complexity more generally.

The second goal of this investigation was to assess whether the typical recruitment of the left anterior temporal lobe during phrasal composition may be modulated by the morphological complexity of the nouns that appear following modifiers. Interestingly, there was a stark absence of anything resembling the typical modification effect in phrases involving opaque compounds, suffixed nouns, or pseudo-suffixed nouns. In contrast, phrasal composition involving transparent compounds elicited increased activation in both of the regions associated with the modification of monomorphemic nouns, though not until approximately 500 ms post-noun onset. These findings are discussed in more detail below.

4.1. *Transparent compounding*

In the pairwise comparisons of transparent compounds, opaque compounds, and monomorphemic nouns, only one significant cluster overlapped with the phrasal composition localizations, capturing increased amplitudes in response to opaque compounds compared to monomorphemic nouns. The functional significance of this effect is not obvious, as a similarly localizing increase in responses to transparent compounds was absent, therefore eliminating an explanation in terms of a general sensitivity to morphological complexity, or a shared morphosyntactic aspect of compound processing. While we could speculate as to the functional nature of this effect, the original motivation for this study was to assess the potential overlap between morphological and phrasal composition. For this reason, we turn to the spatiotemporal effect related to transparent compounding that was identified in the intersection analysis.

The spatiotemporal intersection analysis isolated a near-significant cluster that captured increased responses to transparent compounds relative to both monomorphemic nouns and opaque compounds in the left temporal pole. Although this was found in a secondary analysis, it does match previous results. In general, the finding that transparent compounding elicits increased activity in an anterior sub-region of the left temporal lobe harmonizes with previous results associating this region, more broadly, with conceptual processing and combination (Baron & Osherson, 2011; Blanco-Elorrieta & Pylkkänen, 2016; Boylan et al., 2017; Del Prato & Pylkkänen, 2014), as transparent compounds represent classic instances of conceptual combination. The timing of the observed effect roughly aligns with the result of previous work that has manipulated the semantic transparency of compounds, showing increases in N400-like activation in response to transparent compounds, in acoustic (Koester et al., 2007; MacGregor & Shtyrov, 2013) and visual presentation (Brooks & Cid de Garcia, 2015, though the onset of this effect was earlier in this study). The approximate parallels between the previous effects observed in comparisons of transparent and opaque compounds, to those observed in manipulations of semantic plausibility of incrementally presented constituents (Koester et al., 2007) and the lexical status of the compounds (Fiorentino et al., 2014; MacGregor & Shtyrov, 2013) has been accounted for by positing the presence of a compositional process, that is present only in the recognition of the transparent compounds. In this light, the LATL compounding effect identified here could be interpreted as a correlate of this process. That said, variability in the timing and spatial patterns of the previously identified effects, and the relatively weak nature of the present result, suggest that this compositional process, if it does exist, is not present in a uniform or consistent manner, as would be expected of a general mechanism recruited in every instance of transparent compounding.

Additionally, we do not have sufficient evidence to argue that there is an absolute distinction between compounds that do and do not rely on decomposition and re-combination process, and the related variability may influence the compounding effect identified here. As may be predicted from variants of parallel dual-route accounts, it is possible that only a subset of the transparent compound stimuli rely on a decomposition and re-combination processing strategy for recognition, while others can be recognized via their full-form. Likewise, it could be that a number of our opaque compound stimuli elicit attempts at decomposition and re-combination. Our intent in the design of the stimuli, which followed from the work of Brooks and Cid de Garcia (2015), was to split the compounds into two groups that were maximally likely to differ on this dimension, and replicate the previously identified effects. In a post-hoc analysis of high and low frequency compounds, we failed to identify any significant effects in the comparisons of transparent compounds, opaque compounds, and monomorphemic nouns. We believe that the lower number of trials in each average response, and the corresponding drop in signal-to-noise ratio, contributed to this lack of results. Future work may be able to better assess the nature and the dissociations between high/low frequency transparent and opaque compounds by increasing the number of stimuli that can be classified in each condition so to explicitly test the impact of this factor on responses in the left anterior temporal lobe.

Even in the face of ambiguity regarding its functional significance, having identified a portion of the left anterior temporal lobe that shows a sensitivity to the semantic transparency of compounds, we were able to directly compare the localization of this effect to that associated with phrasal composition, demonstrating dissociation between the correlates in the group-level analysis. While this may be driven by a true distinction between the neural mechanisms involved in each case, we must refrain from interpreting this as strong evidence in support of such a distinction. This is due to the limited spatial resolution of MEG, and the tendency for field spread in MEG source estimates, as well as the equivocal findings from the analyses of overlap in individual subjects. If future work can show a reliable spatial divergence between the localization of effects related to the recognition of low frequency, or novel, transparent compounds and phrasal composition, for example, that could provide compelling evidence for distinct combinatory mechanisms underlying each case.

4.2. *Suffixations*

With regard to suffixation, experiments 1 and 2 each revealed different effects, with the pair of them failing to identify consistent patterns that could be associated with either morpheme-based combination, or a more general sensitivity to morphological complexity. Furthermore, the effects identified were particularly short in duration (approximately 15 ms), and the visualization of the corresponding time-courses did not reveal a compelling dissociation between the conditions. The inconsistencies between the two experiments could be explained by multiple factors. In addition to differences in the MEG systems used to collect the data, the corresponding electromagnetic noise environments, and the pre-processing steps (notably, the application of a 1 Hz high-pass filter in experiment 2), relevant properties of the stimuli also differed across the two studies. In particular, the experiment 1 materials were less frequent and longer than those included in experiment 2. From the perspective of dual-route accounts of complex word recognition, the reduced frequency would seemingly reduce the number of exposures that could contribute to the storage of the lexeme in memory, therefore suggesting a greater reliance on a parsing and re-combination route. In this light, it is possible that the experiment 1 materials more readily elicited responses related to decomposition and re-combination, while the full-form route may have been more heavily relied upon in the recognition of the more frequent experiment 2 materials. If the aSTG effect identified in experiment 1 does in fact reflect processing of the morphemes, or a re-combination process, this frequency difference could explain the failure to find a similar dissociation in this region between the more frequent suffixed and pseudo-suffixed nouns of experiment 2. This is particularly intriguing when considering the results of Fruchter and Marantz (2015), wherein portions of the middle and superior temporal gyri showed response amplitudes that correlated with the derivational family entropy of a stem (a characterization of the statistical distribution of suffixes that a stem occurs with) between 240 and 320 ms post-word onset, suggesting that activity in this spatiotemporal window may index a process akin to the lexical look up of the isolated stem. That being said, this explanation could not account for the inconsistent effects in the aITG fROI, which showed a significant increase unique to suffixed nouns in experiment 2, but no effect in experiment 1. This is puzzling in light of the abovementioned lexical frequency differences, and the characterization of regions in the inferior temporal lobe as underlying decomposition-related processes (Gwilliams, Lewis, & Marantz, 2016; Solomyak & Marantz, 2010), which would suggest that the infrequent nouns in experiment 1 would be more likely to elicit differences related to decomposition in this region.

It is also worth considering further the manner in which suffixation differs from semantically transparent compounding, and how this may relate to the present findings in the left anterior temporal lobe. Theoretical accounts of morphology posit that the relation between stem and suffix represents an instance of function application (see for example, Lieber, 2015; for a brief overview of function application and its relation to semantic composition more broadly, see Pylkkänen, Brennan, & Bemis, 2011), which is in some senses akin to the combination of a verb and its arguments, and can be contrasted with the conceptual combination postulated to take place in compounding, and previously studied instances of adjective-noun composition (e.g., Bemis & Pylkkänen, 2011). More specifically, a suffix can be considered a predicate, whose meaning is incomplete until it takes a stem that satisfies its semantic requirements as its argument (referred to as argument saturation). This characterization serves to highlight the grammatical or functional nature of the meaning of suffixes. Although previous work shows that argument saturation alone does not negate the involvement of the LATL in composition (Westerlund et al., 2015), this was done in the context of verb phrases (e.g., *eats meat*). In such instances, the verbs and their arguments consist of free morphemes that appear in many different contexts, and denote relatively rich concepts (i.e., compare the meaning of *eat* to the meaning of −*ist*). Adopting again the previous characterization of the LATL as being involved in conceptual processing (Baron & Osherson, 2011; Del Prato & Pylkänen, 2014; Patterson et al., 2007), the heavier grammatical nature of suffixes, in combination with possible distinctions in their conceptual representation due to their bound nature, may disqualify the suffixes from being processed in the LATL in the same manner as adjectives, nouns, or verbs, therefore entailing a reliance on another processing system.

4.3. *Phrasal composition involving complex nouns*

Perhaps the most surprising result of this investigation was the absence of a typical phrasal composition effect in the majority of the complex noun conditions. Specifically, there were no significant increases in either fROI, nor across the entire left temporal lobe, in response to modified opaque compounds, suffixed nouns, or pseudo-suffixed nouns, relative to their unmodified counterparts. In the case of suffixed nouns, this was true even when a follow-up analysis was performed using the combined data of both experiments, and fROIs that were specifically defined on the basis of each individual experiment. There are a number of differences across the conditions that may account for and/or contribute to this pattern of results. First, the number of phrases involving an adjectival modifier differed across conditions, with the opaque compound and suffixed stimuli of experiment 1 being paired with more adjectival modifiers (as opposed to nominal modifiers) than both the monomorphemic and transparent compound stimuli. However, there is reason to doubt that the part of speech of the modifier alone caused the null effects. Most notably, the experiment 1 conditions that showed the phrasal composition effect were those that involved fewer adjectival modifiers (monomorphemic nouns and transparent compounds), while the LATL phrasal composition effect was also present in the monomorphemic condition of experiment 2, where all modifiers were adjectives. A second factor that could potentially account for the absence of phrasal composition effects is the bigram frequency of the composed phrases across the types of complex nouns. Specifically, transparent compound phrases, which showed a phrasal composition effect, were more frequent than the remainder of the conditions, while the opaque compounds were the least frequent of the conditions. We thus cannot rule out the possibility that the frequency of phrases interacts with the morphological complexity of the heads to influence the degree to which the LATL is involved in the phrasal composition, and/or the timing of its involvement relative to the onset of the head.

Another explanation for the null effects in the typical 200–300 ms time window may be that the process indexed by this LATL response is bypassed, or otherwise missed, during phrasal composition involving nouns that are morpho-orthographically complex. Related to this, previous results have provided evidence that the 200–300 ms window in which the LATL response typically appears may reflect a stage of opportunistic composition, working with conceptual/semantic representations that have become available by this relatively early point (Ziegler & Pylkkänen, 2016; see Westerlund & Pylkkänen, 2014; Zhang & Pylkkänen, 2015 for related findings). In light of this, if, on average, the presence of morpho-orthographic complexity negates the early availability of semantic representations, one would expect the magnitude of the composition response in this window to be small, and perhaps negligible in condition averages. A similar argument could be made for the presence of simple variability in the timing of access to the semantic representations of the complex noun conditions. That is, if there is considerable variability in the time required to process the stimuli in opaque, suffixed, and pseudo-suffixed conditions prior to their entering phrasal composition, this variability could reduce the likelihood of finding composition-related effects across condition averages. Lastly, given the previous evidence that the left anterior temporal lobe is preferentially recruited in phrasal composition that involves some form of conceptual combination, it could be that only the monomorphemic and transparent compound conditions included enough modifier-noun pairs that met this criterion, which is yet to be precisely defined, to sufficiently drive the effect in condition averages.

Lastly, phrasal composition involving transparent compounds, in contrast to the types of complexity discussed above, did elicit increased levels of activation in both of the regions associated with phrasal composition involving monomorphemic nouns. This appeared first in the anterior superior temporal gyrus between 212 and 258 ms, and then in anterior sections of both the superior and inferior temporal gyri by 530 ms post-onset. The finding that the pair of regions did not show simultaneous increases, as they did during the modification of monomorphemic nouns, until late in the epoch, fits well with both a semantic/conceptual account of the LATL's role in phrasal composition, and the timing of activity that has been associated with within-word composition. In particular, if the activity in the left anterior temporal lobe does indeed index conceptual combination, this would presumably entail that a conceptual representation of the head noun must be arrived at before the response appears in phrasal composition. In the case of transparent compounds, arriving at a conceptual representation of the whole word may require successful within-word composition, posited to occur between 400 and 450 ms, which then feeds into phrasal composition approximately 100 ms later. The early cluster found in the more superior ROI (212–258 ms) may reflect an initial attempt at composition, which is not of sufficient strength or consistency to fully drive both of the regions. If this account is correct, the question of why only the transparent compounds elicit this later response in the left anterior temporal lobe ROIs remains puzzling, as one would have to presume that there must be some point later in time when the representation of semantically opaque compounds, for example, would be available to enter into phrasal composition. One possibility is that it is the presence of a word-internal instance of conceptual combination that enables the “second cycle” of the original processing stream, and thus elicits the late response in the left anterior temporal lobe. Notably, it is possible that, rather than conceptual combination within a word, this second cycle, or delayed onset, could be triggered simply by the simultaneous presentation of two words (i.e., the constituents that make up the compound) that exemplify conceptual combination. The same explanation could apply to the left temporal pole response associated with transparent compounds. In other words, the distinctions (in time and space) between temporal lobe responses associated with semantically transparent compounding, and phrasal composition, could be due to the simultaneous *vs* sequential processing in each context. Interestingly, this can be empirically tested in future work by presenting separate modifiers and nouns on the same screen, to determine whether the 400–450 ms response that was observed in response to transparent compounds can also be elicited by phrases.

4.4. *Additional considerations and future directions*

One factor that has been the topic of much previous work, but which we did not investigate here, concerns the processing of lexicalized versus novel compounds. From the perspective of dual-route accounts of complex word recognition (e.g., Baayen, Dijkstra, & Schreuder, 1997), novel compounds are predicted to rely on a parsing and recombination route, as they do not yet have corresponding lexemes available for look-up. This is especially relevant for the present results, as it could more precisely characterize the stage of processing indexed by the LATL increases present in responses to transparent compounds. If this activity tracks within-word composition, then one would expect a similar increase to be present in this region and time window when individuals are presented with novel compounds, relative to suitable baselines. Extending the interpretation of the present results, if novel compounds appeared following a modifier, we would then also expect the delayed LATL phrasal composition response in these cases. Such a design thus presents itself as a straightforward avenue toward a more precise characterization of phrasal composition and conceptual combination.

Finally, the construction of the stimuli in the present investigation was not done with the intention of accounting or controlling for all fine-grained varieties of semantic composition that exist. One such source of variety that we have been indifferent to in this design is the conceptual relationship that is present in the combination of constituents of transparent compounds (e.g., a doghouse is a house for dogs, while a farmhouse is a house located on a farm). The nature of these relationships has been the subject of numerous frameworks related to conceptual combination (e.g., Gagné & Shoben, 1997), and previous results suggest that they may indeed impact online processing (Gagné, 2002; Schmidtke, Kuperman, Gagné, & Spalding, 2016), with different types of relationships possibly relying on dissociable neural systems (Boylan et al., 2017). Similar to the present work, previous electrophysiological studies of compounding do not yet appear to have taken this dimension of the meaning of transparent compounds into account (see for example, the stimuli used by MacGregor & Shtyrov, 2013), leaving open the question of whether or not it drastically alters the neural underpinnings of the corresponding conceptual combination. This is especially intriguing with regard to the overlap that may exist amongst combinatory operations at different levels. Specifically, although the present study does not find strong evidence in favor of a shared compositional mechanism in transparent compounding and phrasal composition, it could be that such overlap would be found if both instances of composition involve the same relation between combined concepts and/or similarities in the concepts that are modifying/modified (e.g., *doghouse* is a house in which dogs live, and *bird nest* is a nest in which birds live). Recent neuroimaging work (Boylan et al., 2017) indicates that the angular gyrus may be preferentially involved in processing relation-based conceptual combination in novel compounds (e.g., *honey soup* is soup is made from/with honey, or has honey in it), while the left anterior temporal lobe may be involved in more attributive based conceptual combination (e.g., *blimp belly* is a belly that looks like a blimp, interpretable as the concept *belly* being attributed a feature of the modifier, i.e., the shape/size of the concept *blimp*). This could thus provide the basis for a more fine-grained investigation of the overlap in the processing of compound nouns and phrasal composition, sharing or differing in the type of conceptual combination taking place. Whether and/or how the combination of stems and suffixes fit into this schema remains another open question, and one that the neurophysiological effects could potentially speak to (i.e., do novel combinations of stems and particular suffixes pattern in the same manner as specific varieties of compounding?).

5. Conclusion

A complete neurobiological account of language requires the characterization of combinatory operations that take place at many different levels. Our results provide new insight concerning the role of the left anterior temporal lobe in compositional operations within and between words. Replicating previous findings, we demonstrate that sections of the left anterior temporal lobe show increased levels of activation (200–300 ms post-noun onset) in response to monomorphemic nouns that enter into phrasal composition, corroborating the proposal that the LATL is responsible, at least in part, for semantic composition that yields the meaning of the phrase. By finding that semantically transparent compounding (e.g., *tombstone*) also elicits increased activity in a neighboring section of the LATL, we provide evidence to support the notion that this region plays a role in the recognition of compounds whose meaning can be predicted from a combination of their constituents, with responses in this spatiotemporal window possibly underlying the combination of the constituents. No such correlate could be reliably identified for the composition that is proposed to take place between stem and suffix during the visual recognition of suffixed nouns (e.g., *starlet*). Finally, the left anterior temporal lobe response elicited by phrasal composition was absent when the modified noun was an opaque compound, suffixed noun, or pseudo-suffixed noun, suggesting that the underlying process may be influenced by the morphological complexity of the head noun. In contrast, modifier-noun phrases involving head nouns that were semantically transparent compounds (e.g., *granite tombstone*) did elicit the typical composition response, though at a delayed latency (500–600 ms post-onset). We suggest that this delay, and the exclusivity of this late LATL response to the transparent compound condition, may be due to the presence of a within-word process related to conceptual combination, taking place prior to the initialization of the phrasal composition process housed in the left anterior temporal lobe.

Acknowledgements

We extend special thanks to Laura Gwilliams and Esti Blanco-Elorrieta for help during the design of stimuli for Experiment 1. This research was supported by the National Science Foundation grant BCS-1221723 (LP) and grant G1001 from the NYUAD Institute, New York University Abu Dhabi (LP, AM).

Appendix 1

Experiment 1 stimuli.

Monomorphemic	Transparent Compound	Opaque Compound	Suffixed
huge albatross	new baseball	straight beeline	precise accountant
purple amethyst	clean bathroom	leather blackjack	poor actor
scant applause	serene bedroom	formal blacklist	smart arithmetician
fresh bratwurst	warm beefsteak	petty blackmail	cotton bedding
tender broccoli	surprise birthday	master blacksmith	light blockage
sultry burlesque	slim bookshelf	crass blockhead	stiff bondage
black cauldron	vicious bullfight	rear bulkhead	peak brightness
king chaplain	hot campfire	tiny butterfly	abrupt brusqueness
proud chieftain	pale cheekbone	pudgy chatterbox	rude cashier
brown chipmunk	remote coalmine	juice cocktail	swift cruelty
wood clarinet	lush cornfield	metal crowbar	blue droplet
plain couscous	glass doorhandle	serious deadline	wild duckling
fun cribbage	iron drainpipe	secure deadlock	intense duelist
whole croissant	plump earthworm	stale doughnut	adept financier
mini dachshund	magic fairyland	public figurehead	genuine friendship
mild dandruff	brick farmhouse	round flapjack	stable fusion
ground espresso	empty flagpole	quick flashback	clumsy giantess
dense eucalyptus	bare footprint	veteran fullback	heroic guardian
pink flamingo	eerie graveyard	short gangplank	extreme hardship
marble fountain	square hairbrush	crisp gingersnap	discrete hostess
putrid gangrene	limp handshake	eager greenhorn	former informant
stone gargoyle	single honeybee	chief hallmark	foreign knighthood
yellow jaundice	brass keyhole	tight hamstring	sour lemonade
spare kerosene	canvas mailbag	oil heirloom	fierce loyalist
youth lacrosse	steel mousetrap	smug highbrow	holy martyrdom
dead mackerel	strong mouthwash	total hogwash	ripe nectarine
amateur masseuse	frigid northeast	raw horseradish	sweet orangeade
vegan mayonnaise	ornate pillbox	giant jackpot	proper placement
severe migraine	great racehorse	plastic joystick	small planetoid
full moustache	steady rainstorm	smooth kingpin	high priesthood
delicious omelet	busy roadside	loose longhand	mutual reliance
faint perfume	gray sailboat	vital mainstay	fragile seedling
roast pheasant	deep seawater	tart pineapple	rare serpentine
marine plankton	lost shipwreck	simple pitchfork	legal sovereignty
green plaintain	extra shoelace	quiet ragtime	white spheroid
baby porpoise	blank sketchpad	common scapegoat	sudden stardom
bland porridge	crystal snowflake	rapid shorthand	blonde starlet
gas rotisserie	arid southwest	orange snapdragon	honest statement
rough scrimmage	red swimsuit	brave spitfire	moist stuffing
wheat spaghetti	porcelain teacup	tense stalemate	acute tension
crude strumpet	granite tombstone	hard starboard	junior trainee
local synagogue	mint toothpaste	random sweepstake	clever trickster
ancient tortoise	secret videotape	open tailgate	senior trustee
solid turquoise	native wildlife	vile turncoat	clear utterance
brief vignette	flat workbench	shy wallflower	strict warrantor

Experiment 2 stimuli.

Monomorphemic	Suffix	Suffixed	Pseudo-suffixed
dark asphalt	-ade	total blockade	blind arcade
good athlete		pink lemonade	steady parade
extra vitamin		best orangeade	small stockade
soft breeze	-al	grotesque normal	climactic final
purple cabbage		busy signal	narrow canal
modest cottage		cruel betrayal	real jackal
great pyramid		old hymnal	gigantic sandal
usual caveat		cheap rental	eager marshal
bare cement		standard postal	crimson petal
ornate chapel	-ant	humble servant	marital covenant
weak stomach		rival claimant	ugly pageant
white curtain		wrong coolant	marine sergeant
stubborn donkey	-er	quiet dryer	dear brother
major tragedy		deep fryer	quick shower
luscious dessert		low rider	cozy corner
blue dolphin		manual mower	sore shoulder
modern miracle		open sewer	edible flower
tall bamboo		extreme skier	cosmetic counter
thick blanket	-ery	derelict bindery	renal artery
fresh banana		free brewery	cold scullery
bitter harvest		gentle mockery	main gallery
ancient proverb	-ic	special graphic	perverse logic
warm sausage		female alcoholic	sheer panic
black trumpet		late classic	new tactic
grand lounge		single metric	plain antic
giant magnet		true psychic	broad topic
short paddle	-ion	pure fusion	entire mission
evil penguin		slow erosion	half billion
firm pillow		minor faction	meager ration
safe refuge		recent edition	loyal legion
stale tobacco		swift action	spare trillion
flat horizon		certain tension	rich lotion
raw shrimp	-let	tiny droplet	whole gauntlet
ripe tomato	-let	simple eyelet	empty goblet
utter despair		little piglet	bright scarlet
sweet vinegar	-ment	full payment	posh department
foreign embassy		wet pavement	brittle parchment
fantastic voyage		common ailment	mere figment
fat candle	-or	serious error	young pastor
shrill whistle		net debtor	gray pallor
vast corpus		avid sailor	local tailor
hot potato		remote sensor	false rumor
heavy perfume		holy savior	strange clamor
legal jargon		mobile vendor	high tenor
brown rabbit	-ure	sudden erasure	large culture
big whisky		dismal failure	nice feature
dead parrot		secret rapture	green manure
thin stripe		possible closure	rough texture
wide terrace		odd mixture	bold venture
strong tornado		mild seizure	romantic overture

Additional details.

One stem was repeated in the experiment 1 suffixed noun materials (*starlet*, *stardom*). No stems were repeated in the experiment 2 materials. Pseudo-suffixed nouns were filtered through the following exclusion criteria: no pseudo-stems were inflected nouns (e.g., sat-an, peas-ant), no pseudo-stems were abbreviations (e.g., lab-or), no pseudo-stems contained a sequence of vowels at the morphological boundary (e.g., do-or, care-er), and no pseudo-stems were repeated. Each pseudo-suffixed noun was matched to a corresponding suffixed noun in experiment 2 (e.g., *piglet/scarlet*), ensuring that each (pseudo-)suffix appeared the same number of times across the two conditions.

The two sets of suffixed nouns differed from each other in a number of ways that could impact online processing of them, and these differences motivated the separate analyses of the corresponding data. First, the experiment 1 materials were longer, and less frequent. Second, the distribution of word classes of the stems differed across the two experiments, with a higher proportion of stems tagged with the verb part-of-speech label in the English Lexicon Project in experiment 2 (39/50 items, approximately 78%) than experiment 1 (25/46 items, approximately 54%). Lastly, the experiment 2 suffixed nouns were processed in the presence of pseudo-suffixed nouns, which could alter the strategy adopted by a system that may rely on decomposition/re-combination. Notably, previous behavioral work suggests that during lexical decision tasks, simply having non-word foil stimuli that contain real-word stems can influence the type of processing that takes place (Taft, 2004), making it possible that the presence of pseudo-suffixed stimuli in the second experiment could have impacted the way in which the suffixed nouns were processed.

One constituent (*room*) was repeated as a head within the set of transparent compounds. Four constituents were repeated within the set of opaque compounds. The constituent *black* appeared as a modifier in four compounds (*blackjack, blacklist, blackmail, blacksmith*), *dead* appeared as the modifier in two compounds (*deadline, deadlock*), *back* appeared as the head in two compounds (*flashback, fullback*) and *hand* appeared as the head in two compounds (*longhand, shorthand*). During stimulus selection, we did not control for factors such as the conceptual relationship between the compound constituents in the transparent condition, nor the word class of the constituents that make up the compounds in both the opaque and transparent conditions. Post-hoc inspection of the stimulus sets revealed a heterogeneous sample of conceptual relationships in the set of transparent compounds (e.g., *snowflake:* a flake of snow; *campfire:* a fire made while camping; *mousetrap:* a trap intended for mice). Post-hoc inspection of the word class of constituents that made up the transparent and opaque compounds revealed that while 33 compounds in the transparent condition could be straightforwardly considered noun–noun compounds, this was only considered true of 20 compounds in the opaque condition. The remainder of the compounds in each condition consisted of adjective-noun (e.g., *wildlife*, *highbrow*), noun-verb (e.g., *mouthwash*, *hogwash*), verb-noun (e.g., *swimsuit*, *turncoat*), and ambiguous cases (e.g., *graveyard*, *hallmark*).

Bigram frequencies of the phrases (displayed in Tables 2 and 3) were not controlled for in either experiment. Pairwise t-tests and Wilcoxon-Mann-Whitney tests did not reveal any significant differences ($p > .05$) between the conditions, however a near-significant difference was found in the comparison of opaque compound and monomorphemic conditions in experiment 1 (Wilcoxon-Mann-Whitney $U = 1289$, $n_1 = n_2 = 46$, $p < .10$ two-tailed). Inspection of the transparent compound materials revealed that three phrases (*new baseball*, *surprise birthday*, *native wildlife*) were much more frequent than all other items in experiment 1 (bigram frequencies of 96, 49, and 50, respectively; all other bigram frequencies < 43), contributing to greater variability in this characteristic. Since the primary analyses utilized responses that were averaged across all items (after trial rejection), we did not remove these items, as they likely had only a minor influence on the evoked responses. The experiment 1 conditions were comparable on the number of phrases that were not listed in COCA (monomorphemic nouns: 8; transparent compounds: 14, opaque compounds: 11, suffixed nouns: 13). Relative to the experiment 1 materials, the phrases used in experiment 2 were considerably more frequent, and showed greater transition probabilities (compare Tables 2 and 3). In experiment 2, all phrases appeared at least once in COCA.

The English Lexicon Project listed all of the selected modifiers as monomorphemic. In experiment 1, modifiers consisted of both adjectives and nouns, and the number of each differed across the conditions. In particular, adjectives made up 34 of the modifiers on monomorphemic nouns, 32 on transparent compounds, 38 on opaque compounds, and 44 on suffixed nouns. In experiment 2, all modifiers were adjectives except for one (*half* in *half trillion*).

REFERENCES

Adachi, Y., Shimogawara, M., Higuchi, M., Haruta, Y., & Ochiai, M. (2001). Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. *IEEE Transactions on Applied Superconductivity, 11*(1), 669–672.

Aronoff, M. (1994). *Morphology by itself: Stems and inflectional classes.* MIT Press.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singular and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language, 37*, 94–117.

Bai, C., Bornkessel-Schlesewsky, I., Wang, L., Hung, Y. C., Schlesewsky, M., & Burkhardt, P. (2008). Semantic composition engenders an N400: Evidence from Chinese compounds. *Neuroreport, 19*(6), 695–699.

Balota, D., Yap, M. J., Cortese, M. J., Hutchison, K.a., Kessler, B., Loftis, B., et al. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445–459.

Baron, S. G., & Osherson, D. (2011). Evidence for conceptual combination in the left anterior temporal lobe. *Neuroimage, 55*(4), 1847–1852.

Bertram, R., Laine, M., & Karvinen, K. (1999). The interplay of word formation type, affixal homonymy, and productivity in lexical processing: Evidence from a morphologically rich language. *Journal of psycholinguistic research, 28*(3), 213–226.

Bemis, D. K., & Pylkkänen, L. (2011). Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *The Journal of Neuroscience, 31*(8), 2801–2814.

Bemis, D. K., & Pylkkänen, L. (2012). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex, 23*(8), 1859–1873.

Bemis, D. K., & Pylkkänen, L. (2013a). Flexible composition: MEG evidence for the deployment of basic combinatorial linguistic mechanisms in response to task demands. *Plos One, 8*(9).

Bemis, D. K., & Pylkkänen, L. (2013b). Combination across domains: An MEG investigation into the relationship between mathematical, pictorial, and linguistic processing. *Frontiers in Psychology, 3*, 1–20.

Blanco-Elorrieta, E., & Pylkkänen, L. (2016). Composition of complex numbers: Delineating the computational role of the left anterior temporal lobe. *Neuroimage, 124*, 194–203.

Boylan, C., Trueswell, J. C., & Thompson-Schill, S. L. (2017). Relational vs. attributive interpretation of nominal compounds differentially engages angular gyrus and anterior temporal lobe. *Brain and language, 169*, 8–21.

Bozic, M., Fonteneau, E., Su, L., & Marslen-Wilson, W. D. (2015). Grammatical analysis as a distributed neurobiological function. *Human Brain Mapping, 36*, 1190–1201.

Brennan, J., & Pylkkänen, L. (2012). The time-course and spatial distribution of brain activity associated with sentence processing. *Neuroimage, 60*(2), 1139–1148.

Brooks, T. L., & Cid de Garcia, D. (2015). Evidence for morphological composition in compound words using MEG. *Frontiers in Human Neuroscience, 9*(APR), 215.

Chadwick, M. J., Anjum, R. S., Kumaran, D., Schacter, D. L., Spiers, H. J., & Hassabis, D. (2016). Semantic representations in the temporal pole predict false memories. *Proceedings of the National Academy of Sciences, 113*(36), 10180–10185.

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic statistical parametric Mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron, 26*(1), 55–67.

Davies, M. (2008). *The Corpus of Contemporary American English (COCA): 520 million words, 1990-present*. Available online at http://corpus.byu.edu/coca/.

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature neuroscience, 19*(1), 158.

Del Prato, P., & Pylkkänen, L. (2014). MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. *Frontiers in Psychology, 5*, 524. June.

Fedorenko, E., Scott, T. L., Brunner, P., Coon, W. G., Pritchett, B., Schalk, G., et al. (2016). Neural correlate of the construction of sentence meaning. *Proceedings of the National Academy of Sciences, 113*(41), E6256–E6262.

Fiorentino, R., Naito-Billen, Y., Bost, J., & Fund-Reznicek, E. (2014). Electrophysiological evidence for the morpheme-based combinatoric processing of English compounds. *Cognitive Neuropsychology, 31*(1–2), 123–146.

Frankland, S. M., & Greene, J. D. (2015). An architecture for encoding sentence meaning in left mid-superior temporal cortex. *Proceedings of the National Academy of Sciences, 112*(37), 11732–11737.

Friederici, A. D. (2012). The cortical language circuit: From auditory perception to sentence comprehension. *Trends in Cognitive Sciences, 16*(5), 262–268.

Friederici, A. D., Opitz, B., & Von Cramon, D. Y. (2000). Segregating semantic and syntactic aspects of processing in the human brain: An fMRI investigation of different word types. *Cerebral Cortex, 10*(7), 698–705.

Fruchter, J., & Marantz, A. (2015). Decomposition, lookup, and recombination: MEG evidence for the full decomposition model of complex visual word recognition. *Brain and Language, 143*, 81–96.

Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier–noun combinations. *Journal of Experimental Psychology Learning, Memory, and Cognition, 23*(1), 71.

Gagné, C. L. (2002). The competition-among-relations-in-nominals theory of conceptual combination: Implications for stimulus class formation and class expansion. *Journal of the Experimental Analysis of Behavior, 78*(3), 551.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience, 7*, 1–13.

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., et al. (2014). MNE software for processing MEG and EEG data. *Neuroimage, 86*, 446–460.

Gwilliams, L., Lewis, G. A., & Marantz, A. (2016). Functional characterization of letter-specific responses in time, space and current polarity using magnetoencephalography. *Neuroimage, 132*, 320–333.

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Reviews Neuroscience, 37*, 347–362.

Häikiö, T., Bertram, R., & Hyönä, J. (2011). The development of whole-word representations in compound word processing: Evidence from eye fixation patterns of elementary school children. *Applied Psycholinguistics, 32*(3), 533–551.

Halle, M., & Marantz, A. (1993). Distributed morphology and the pieces of inflection. In K. Hale, & S. J. Keyser (Eds.), *The view from building 20* (pp. 111–176). Cambridge: MIT Press.

Hanna, J., & Pulvermüller, F. (2014). Neurophysiological evidence for whole form retrieval of complex derived words: A mismatch negativity study. *Frontiers in Human Neuroscience, 8*(886).

Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience, 18*(4), 665–679.

Humphries, C., Love, T., Swinney, D., & Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping, 26*(2), 128–138.

Koester, D., Gunter, T. C., & Wagner, S. (2007). The morphosyntactic decomposition and semantic composition of German compound words investigated by ERPs. *Brain and Language, 102*, 64–79.

Koester, D., Holle, H., & Gunter, T. C. (2009). Electrophysiological evidence for incremental lexical–semantic integration in auditory compound comprehension. *Neuropsychologia, 47*, 1854–1864.

Kuperman, V., Bertram, R., & Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes, 23*(7–8), 1089–1132.

Kuperman, V., Schreuder, R., Bertram, R., & Baayen, R. H. (2009). Reading polymorphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology Human Perception and Performance, 35*(3), 876.

Leminen, A., Leminen, M., Kujala, T., & Shtyrov, Y. (2013). Neural dynamics of inflectional and derivational morphology processing in the human brain. *Cortex, 49*, 2758–2771. https://doi.org/10.1016/j.cortex.2013.08.007.

Lieber, R. (2015). *Introducing morphology*. Cambridge University Press.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers, 28*(2), 203–208.

Mankin, J. L., Thompson, C., Branigan, H. P., & Simner, J. (2016). Processing compound words: Evidence from synaesthesia. *Cognition, 150*, 1–9.

Marantz, A. (1997). No escape from Syntax: Don't try morphological analysis in the privacy of your own Lexicon. *University of Pennsylvania Working Papers in Linguistics, 4*(2), 201–225.

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

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., & Salamon, G. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience, 5*(4), 467–479.

MacGregor, L. J., & Shtyrov, Y. (2013). Multiple routes for compound word processing in the brain: Evidence from EEG. *Brain and Language, 126*(2), 217–229.

Nelson, M. J., El Karoui, I., Giber, K., Yang, X., Cohen, L., Koopman, H., et al. (2017). Neurophysiological dynamics of phrase-structure building during sentence processing. *Proceedings of the National Academy of Sciences, 114*, E3669–E3678.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113.

Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences of the United States of America, 108*(6), 2522–2527.

Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience, 8*(12), 976–987.

Poortman, E. B., & Pylkkänen, L. (2016). Adjective conjunction as a window into the LATL's contribution to conceptual combination. *Brain and language, 160*, 50–60.

Pylkkänen, L., Bemis, D. K., & Elorrieta, E. B. (2014). Building phrases in language production: An MEG study of simple composition. *Cognition, 133*(2), 371–384.

Pylkkänen, L., Brennan, J., & Bemis, D. K. (2011). Grounding the cognitive neuroscience of semantics in linguistic theory. *Language and Cognitive Processes, 26*(9), 1317–1337.

Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes, 23*(7–8), 942–971.

Rogalsky, C., & Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebral Cortex, 19*(4), 786–796.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 131–154). Hillsdale, NJ: Erlbaum.

Schmidtke, D., Kuperman, V., Gagné, C. L., & Spalding, T. L. (2016). Competition between conceptual relations affects compound recognition: The role of entropy. *Psychonomic bulletin & review, 23*(2), 556–570.

Stites, M. C., Federmeier, K. D., & Christianson, K. (2016). Do morphemes matter when reading compound words with transposed Letters? Evidence from eye-tracking and event-related potentials. *Language, Cognition and Neuroscience, 31*(10), 1299–1319.

Solomyak, O., & Marantz, A. (2010). Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience, 22*(9), 2042–2057.

Stowe, L. A., Broere, C. A. J., Paans, A. M. J., Wijers, A. A., Mulder, G., Vaalburg, W., et al. (1998). Localizing components of a complex task: Sentence processing and working memory. *Neuroreport, 9*(13), 2995–2999.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *The Quarterly Journal of Experimental Psychology Section A, 57*(4), 745–765.

Tanner, D., Morgan-Short, K., & Luck, S. J. (2015). How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology, 52*(8), 997–1009.

Van Essen, D. C. (2005). A Population-Average, Landmark- and Surface-based (PALS) atlas of human cerebral cortex. *Neuroimage, 28*(3), 635–662.

Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience, 14*(4), 550–560.

Westerlund, M., Kastner, I., Al Kaabi, M., & Pylkkänen, L. (2015). The LATL as locus of composition: MEG evidence from English and Arabic. *Brain and Language, 141*, 124–134.

Westerlund, M., & Pylkkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia, 57*(1), 59–70.

Williams, A., Reddigari, S., & Pylkkänen, L. (2017). Early sensitivity of left perisylvian cortex to relationality in nouns and verbs. *Neuropsychologia, 100*, 131–143.

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *Neuroimage, 25*(3), 1002–1015.

Zhang, L., & Pylkkänen, L. (2015). The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. *Neuroimage, 111*, 228–240.

Ziegler, J., & Pylkkänen, L. (2016). Scalar adjectives and the temporal unfolding of semantic composition: An MEG investigation. *Neuropsychologia, 89*, 161–171.