



Composing lexical versus functional adjectives: Evidence for uniformity in the left temporal lobe

Linmin Zhang^{1,2}  · Liina Pykkänen^{1,3,4}

Published online: 24 April 2018
© Psychonomic Society, Inc. 2018

Abstract

Featural information (e.g., color or shape) allows interlocutors to focus their attention on the specific items under discussion from the vast set of possibilities in the environment. Intriguingly, when they are used to modify and restrict nouns, adjectives can either carry featural information themselves (e.g., *green car*) or retrieve featural information from the context (e.g., somebody points at a car and claims that she has the *same car* or a *different car*). Do the processing of *same/different car* and *green car* share neural correlates? For the composition of nouns with feature-carrying adjectives, prior work revealed early compositional effects (roughly 200 ms after noun onset) in the left anterior temporal lobe. However, although we know that such effects do not extend to cases of numeral quantification, which add no conceptual features to the noun (e.g., *two boats*), we do not know whether they extend to functional adjectives that themselves introduce no features, but instead reference features in the context. To address this question, we measured magnetoencephalography (MEG) during the processing of five types of noun phrases (NPs): *same* NPs (e.g., *same star*), *different* NPs (e.g., *different star*), color NPs (e.g., *green star*), comparative NPs (e.g., *larger star*), and *another* NPs (e.g., *another star*). Our main finding was that between 185 to 240 ms after noun onset, *same* and *different* NPs patterned with the color NPs in their elicited left temporal lobe activity, and *same* NPs even trended toward higher amplitudes than the color NPs. This shows that the mechanism driving combinatory effects in the left temporal cortex does not require the input words to directly name conceptual features, as long as the words reference featural information in the context, and that overlapping neural correlates underlie the composition of featural information from both linguistic and nonlinguistic sources.

Keywords Language comprehension · Semantics · Neuropsychology · Psycholinguistics

Composing featural information via lexical versus functional adjectives

Lexical adjectives, such as *green* and *silky*, carry featural information that often plays a specifying and restrictive role in human communication, helping interlocutors focus their attention on the specific items under discussion from the vast

set of possibilities in the environment. Intriguingly, some functional adjectives, such as *same* and *different*, despite lacking featural information themselves, can nevertheless play a similar role. For instance, if someone points at a dark green BMW sports car, claiming that she has the same car in a different color, we can easily infer that the speaker also has a BMW sports car, but that her car is not dark green. This example shows that within a given context, using *same* or *different* can not only convey highly specific featural information, but with an impressive efficiency. The power of human language lies precisely in this efficiency, but the underlying neural mechanisms are still understudied. The present study built upon our extant understanding of the processing of lexical adjectives to further investigate processing of the terms *same* and *different*.

Studies using magnetoencephalography (MEG) have revealed that, as compared to unstructured control stimuli, minimal noun phrases (NPs) containing a modifier—for instance, *red boat*—lead to an increased signal in the left anterior

✉ Linmin Zhang
linmin.zhang@nyu.edu

¹ Department of Linguistics, New York University, New York, NY, USA

² New York University Shanghai, 1555 Century Avenue, Shanghai 200122, China

³ Department of Psychology, New York University, New York, NY, USA

⁴ New York University Abu Dhabi Institute, Abu Dhabi, United Arab Emirates

temporal lobe (LATL), peaking around 200–250 ms after onset of the head noun (Bemis & Pykkänen, 2011, 2013a, 2013b; Del Prato & Pykkänen, 2014; Westerlund, Kastner, Al Kaabi, & Pykkänen, 2015; Westerlund & Pykkänen, 2014; Zhang & Pykkänen, 2015; Ziegler & Pykkänen, 2016). Moreover, higher activity amplitudes are elicited by a noun (say *dish*) when it is being integrated with a specific modifier such as *lamb* (yielding *lamb dish*) than when it is integrated with the more general *meat* (yielding *meat dish*) (Zhang & Pykkänen, 2015). These findings reconcile previous results from hemodynamic, electromagnetic and patient studies, which have implicated this brain region as a center for both conceptual knowledge (Hodges, Patterson, Oxbury, & Funnell, 1992; Mummery et al., 2000; Mummery, Patterson, et al., 1999a; Mummery, Shallice, & Price, 1999b) and linguistic composition (Bemis & Pykkänen, 2011; Binney, Embleton, Jefferies, Parker, & Lambon Ralph, 2010; Brennan & Pykkänen, 2012; Dronkers, Wiilkins, Vanvalin, Redfern, & Jaeger, 1994; Humphries, Binder, Medler, & Liebenthal, 2006; Humphries, Love, Swinney, & Hickok, 2005; Humphries, Willard, Buchsbaum, & Hickok, 2001; Mazoyer et al., 1993; Pallier, Devauchelle, & Dehaene, 2011; Stowe et al., 1998; Vandenberghe, Nobre, & Price, 2002). Together, our emerging understanding for the neural basis of noun-phrase-level compositional semantics suggests early effects of featural information integration that can be modulated by featural specificity. Crucially, such compositional effects do not extend to cases of numeral quantification (e.g., *two boats*), which add no conceptual features to the noun (Blanco-Elorrieta & Pykkänen, 2016; Del Prato & Pykkänen, 2014).

As is already shown in our above-mentioned example, the use of *same* and *different* conveys the homogeneity or heterogeneity of featural information in a given context. This view is supported by formal semantics research (Zhang, 2016) and animal studies on the cognition of identification (Martinho & Kacelnik, 2016). Given that *same* and *different* carry no featural information themselves, their actual interpretation has to depend on featural information from potentially nonlinguistic contexts. Thus, in processing phrases like *green car* and *same car*, are shared neural mechanisms recruited for feature integration?

If the brain region performing noun-phrase-level compositional semantics integrates featural information from various sources, linguistic or nonlinguistic, to rapidly compose and form a holistic semantic representation in interpreting an NP, then it should process *same* and *different* as disguised lexical adjectives and integrate context-dependent feature information transferred via the use of these functional adjectives. Furthermore, given that *same* is efficient in transferring multidimensional feature information, it could plausibly elicit even larger compositional effects than lexical adjectives like *green*, which carry feature information on a single dimension (e.g., color).

Alternatively, if the early combinatory process (about 200–250 ms) requires the composing words to directly name concepts, the functional adjectives *same* and *different* should not activate the composition-related brain region as much as lexical adjectives, due to lacking featural information in their word meaning.

To distinguish between these hypotheses, we adapted the canonical adjective-noun paradigm to include a minimal pictorial context. Specifically, participants saw a display of three objects (with the middle one concealed) before reading an NP, and then upon seeing a postverbal display, participants indicated whether a revealed middle item matched the verbal description. For example, if the first display contained a green striped star and the NP was *same star*, a matched final display should show that the previously hidden item was also a green striped star. With this minimal adjustment, we aimed to model the real-world situation in which functional adjectives retrieve features from context, while maintaining full control of the number and content of these features.

Possible correlates of relationality and retrieval

In addition to addressing the contrast between lexical and functional adjectives in combinatory semantic processing, our design also allowed us to explore issues pertaining relational processing and semantic retrieval. Specifically, we tested whether the processing of comparative NPs activates the angular gyrus (AG) and/or the adjacent supramarginal gyrus (SMG)—that is, areas hypothesized as being sensitive to relationality (Ardila, Concha, & Rosselli, 2000; Barbieri, Aggujaro, Molteni, & Luzzatti, 2015; Damasio et al., 2001; de Zubizaray, Hansen, & McMahon, 2013; Humphries, Binder, Medler, & Liebenthal, 2007; Lewis, Poeppel, & Murphy, 2015; Thompson et al., 2007; Williams, Reddigari, & Pykkänen, 2017)—and whether the processing of more context-dependent conditions (i.e., *same* and *different* NPs) activates the left inferior frontal gyrus (LIFG), a region implicated in dependency or retrieval effects (Ben-Shachar, Hendler, Kahn, Ben-Bashat, & Grodzinsky, 2003; Ben-Shachar, Palti, & Grodzinsky, 2004; Caplan, Alpert, Waters, & Olivieri, 2000; Caplan, Stanczak, & Waters, 2008; Constable et al., 2004; Just, Carpenter, Keller, Eddy, & Thurlborn, 1996; Keller, Carpenter, & Just, 2001; Leiken, McElree, & Pykkänen, 2015; Leiken & Pykkänen, 2014; Rogalsky, Matchin, & Hickok, 2008; Stromswold, Caplan, Alpert, & Rauch, 1996).

The AG has also been proposed in neuroanatomical models as a center of semantic processing (Binder & Desai, 2011; Bonner, Peelle, Cook, & Grossman, 2013; Lau, Almeida, Hines, & Poeppel, 2009; Noonan, Jefferies, Visser, & Lambon Ralph, 2013; Patterson, Nestor, & Rogers, 2007;

Price, Bonner, Peelle, & Grossman, 2015; Price, 2010; Seghier, 2013; Seghier, Fagan, & Price, 2011), based both on hemodynamic (Damasio et al., 2001; Graves, Binder, Desai, Conant, & Seidenberg, 2010; Humphries et al., 2007; Newman, Pancheva, Ozawa, Neville, & Ullman, 2001; Pallier et al., 2011; Price, 2010) and patient studies (Ardila et al., 2000; Hart & Gordon, 1990). More specifically, the AG along with the adjacent SMG has been implicated in the processing of thematic relationality (Barbieri et al., 2015; Damasio et al., 2001; de Zubicaray et al., 2013; Humphries et al., 2007; Lewis et al., 2015; Thompson et al., 2007). For example, the MEG study of Lewis et al. (2015) showed that processing the thematic relation between *king* and *castle* robustly activates the AG. Moreover, the patient study of Ardila et al. (2000) showed that the AG is implicated in the processing of more abstract and less eventive relations, such as the meaning of comparatives (e.g., *taller*, *shorter*, *older*, *younger*). The PET study of Damasio et al. (2001) showed that the processing of action names involving transitive verbs elicited more activity in the left SMG than those involving intransitive verbs, and a similarly localized effect was also reported for the naming of spatial relations in comparison to concrete entities. A recent MEG study by Williams et al. (2017) employed eventive and noneventive nouns of high and low relationality (e.g., noneventive nouns: high-relational *daughter* vs. low-relational *girl*; eventive nouns: high-relational *slaughter* vs. low-relational *laughter*) and revealed that the AG is sensitive to relationality but not to eventivity. Based on this understanding of the function of the AG and SMG, we also tested whether and when the processing of comparative NPs elicits relationality effects in the AG/SMG, and furthermore, whether the interpretation of *same* and *different* involves comparison in a similar way and thus activates the AG/SMG as much as comparatives do.

In addition to relational processing, our design also required retrieval of information depicted in the picture context. A relevant finding to this is that the LIFG has been implicated for retrieval mechanisms in sentence processing, specifically for the comprehension of syntactic dependencies both in hemodynamic (Ben-Shachar et al., 2003; Ben-Shachar et al., 2004; Constable et al., 2004; Caplan et al., 2000; Caplan et al., 2008; Just et al., 1996; Keller et al., 2001; Rogalsky et al., 2008; Stromswold et al., 1996) and in MEG studies (Leiken et al., 2015; Leiken & Pylkkänen, 2014). Among the NPs used in this experiment, the interpretation of *same* and *different* NPs should involve the retrieval of semantic information from a pictorial context, and thus potentially increased LIFG activity. Therefore, we also tested whether the processing of *same* and *different* NPs elicits retrieval effects in the LIFG.

In sum, our study centered on investigating the nature of possible input items to feature integration in the left temporal lobe while also assessing the roles of the AG/SMG and LIFG

in task components engaging relational processing and retrieval.

Method

Participants

A group of 23 right-handed native English speakers participated in the study. All had normal or corrected-to-normal vision and gave informed consent. Two participants were excluded from the MEG analyses, one due to excessive noise during the recording, and another for being an outlier in terms of accuracy in the behavioral task (2.69 *SDs* lower than the group mean, whereas the other 21 participants were at most 1.66 *SDs* away from the mean). Thus, 21 participants were included in MEG data analyses (15 females and 6 males; average age = 24.29 years, *SD* = 8.13 years). All data were collected at the Neuroscience of Language Lab (NeLLab) at New York University Abu Dhabi (in NYUAD Saadiyat Campus, UAE).

Experimental design and stimuli

A principal challenge for the study of the functional vocabulary is that each functional item is potentially representationally distinct—in important ways—from each of the other items. This challenge clearly holds for the present study: We do not have a group of examples that share the properties of *same* or *different*, which are entirely related to a context. Thus, although we believe that *same* and *different* can offer an important window into the function of left temporal cortex in combinatorial processing, it is not possible to construct a group of stimuli belonging to the relevant class. To keep the variability of the control conditions matched to the test conditions, they too must be similarly limited to a small set of items. Although the functional vocabulary is small, it is crucial to language, and consequently, studying the individual items is critical for a full understanding of the language system, despite methodological challenges.

In all, our experiment employed five combinatorial conditions—*same*, *different*, color, comparative, and *another* NPs—as well as a sixth noncombinatorial condition—a consonant string NP (see Table 1). In total, there were eight distinct items for the modifier position: *same*, *different*, *green*, *brown*, *larger*, *smaller*, *another*, and *ykdunjz* (i.e., two distinct items for each of the crucial categories: functional adjectives carrying no featural information, adjectives involving size-related features, and adjectives carrying color features). The choice of *green* and *brown* was motivated mainly by their similarity in length and frequency to other words. Each of these eight items was combined with 15 head nouns (*arrow*, *bell*, *boat*, *bottle*, *cross*, *glass*, *heart*, *house*, *lamp*, *plane*, *shirt*,

Table 1 Experimental design and predictions

Condition	Example	Prediction of the Hypothesis in Which Input Words to Composition in the Left Temporal Lobe Must Be Labels of Concepts	Prediction of the Hypothesis in Which Input Words to Composition in the Left Temporal Lobe Need Not Label Concepts as Long as They Reference Concepts
Same NP	<i>same star</i>	No compositional effects in the left temporal lobe	Compositional effects in the left temporal lobe
Different NP	<i>different star</i>		
Color NP	<i>green star</i> <i>brown star</i>	Compositional effects in the left temporal lobe (Bemis & Pylkkänen, 2011, 2013a, 2013b; Westerlund et al., 2015; Westerlund & Pylkkänen, 2014; Zhang & Pylkkänen, 2015; Ziegler & Pylkkänen, 2016)	
Comparative NP	<i>larger star</i> <i>smaller star</i>	Compositional effects for gradable adjectives? (Without information of comparison class: No; Ziegler & Pylkkänen, 2016; With information of comparison class: Yes; Poortman & Pylkkänen, 2016)	
Another NP	<i>another star</i>	Baseline (Blanco-Elorrieta & Pylkkänen, 2016; Del Prato & Pylkkänen, 2014)	
Consonant NP	<i>ykdmpz star</i>	Potential baseline	

shoe, square, star, and vase), forming 120 distinct phrases. Each phrase was used four times in the whole experiment, so that there were 480 critical trials, varying in whether the head noun corresponded to the shape shown on the left or the right in the preverbal picture and whether the correct answer in the behavioral task should be “match” or “mismatch.”

The condition with a meaningless consonant string has served as a noncombinatory baseline in previous MEG investigations of phrasal-level composition (e.g., Bemis & Pylkkänen, 2011). However, in the present study it was possible that inserting a meaningless string between two meaningful trial parts—the pictorial context and the head noun—might awkwardly interrupt processing, and thus we were aware that this condition might not work as a baseline as straightforwardly as in studies without a context. Therefore, we included the *another* NP as another baseline: *another* contributes no featural information, and even though it provides quantity-related information, we already know that quantificational phrases do not elicit LATL composition effects (Blanco-Elorrieta & Pylkkänen, 2016; Del Prato & Pylkkänen, 2014). Moreover, the results from previous MEG studies (Poortman & Pylkkänen, 2016; Ziegler & Pylkkänen, 2016) suggested that size-related adjectives elicit LATL compositional effects only when the comparison class of the noun (e.g., whether a man is classified as tall when compared to a class of professional basketball players or a class of average men) can be inferred from context. Thus, here we included comparative NPs (e.g., *larger star, smaller boat*) in order to test this issue again.

As is shown in Fig. 1, on each trial, following the fixation, first a pictorial context was presented, consisting of a row of three items with the middle one concealed, and then an NP presented word by word. Afterward, when a postverbal picture appeared with all items revealed, participants made a judgment to indicate whether the NP matched the previously concealed item. The pictorial context stayed on the screen for

600 ms, and the fixation and each word stayed on the screen for 300 ms, with 300-ms blank spaces between adjacent items.

The felicitous use of *same, different, comparative, and another* NPs in our trial structure required that one of the two visible items in the pictorial context depicted the shape described by the head noun. To satisfy this requirement and to equate the predictability of the following head noun, the two initially visible items were of different shapes, and the head noun always corresponded to one of these two shapes. In the whole experiment, the shapes varied in color (green vs. brown), pattern (striped vs. nonstriped), and size (small vs. middle vs. large). To balance the predictability, visible items in the context were all of the middle size.

To enforce participants’ attention to the preverbal pictures, we added 48 filler trials with the same kind of preverbal pictures. But then, instead of reading a two-word NP and doing the “phrase–picture” matching task, participants read the short question “have you seen this item?” and judged whether the item shown in the postverbal picture had appeared in the preverbal picture. Thus, altogether, there were 528 trials. The answers “match (yes)” versus “mismatch (no)” were evenly distributed among trials. Pictures used for the postverbal task always stayed on the screen until participants pressed a button to respond.

Participants were asked not to blink during the presentation of stimuli. Before each trial, we added a line of instruction “Press when you are ready. DO NOT BLINK!” to remind them and give them some time to relax their eyes. This instruction stayed on the screen until participants pressed any button to initiate the next trial. The stimuli were presented by Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), in lowercase letters with 30-point Courier font on a gray background. The 528 trials were divided into 12 blocks, each containing 40 critical trials, quasi-evenly distributed among the types

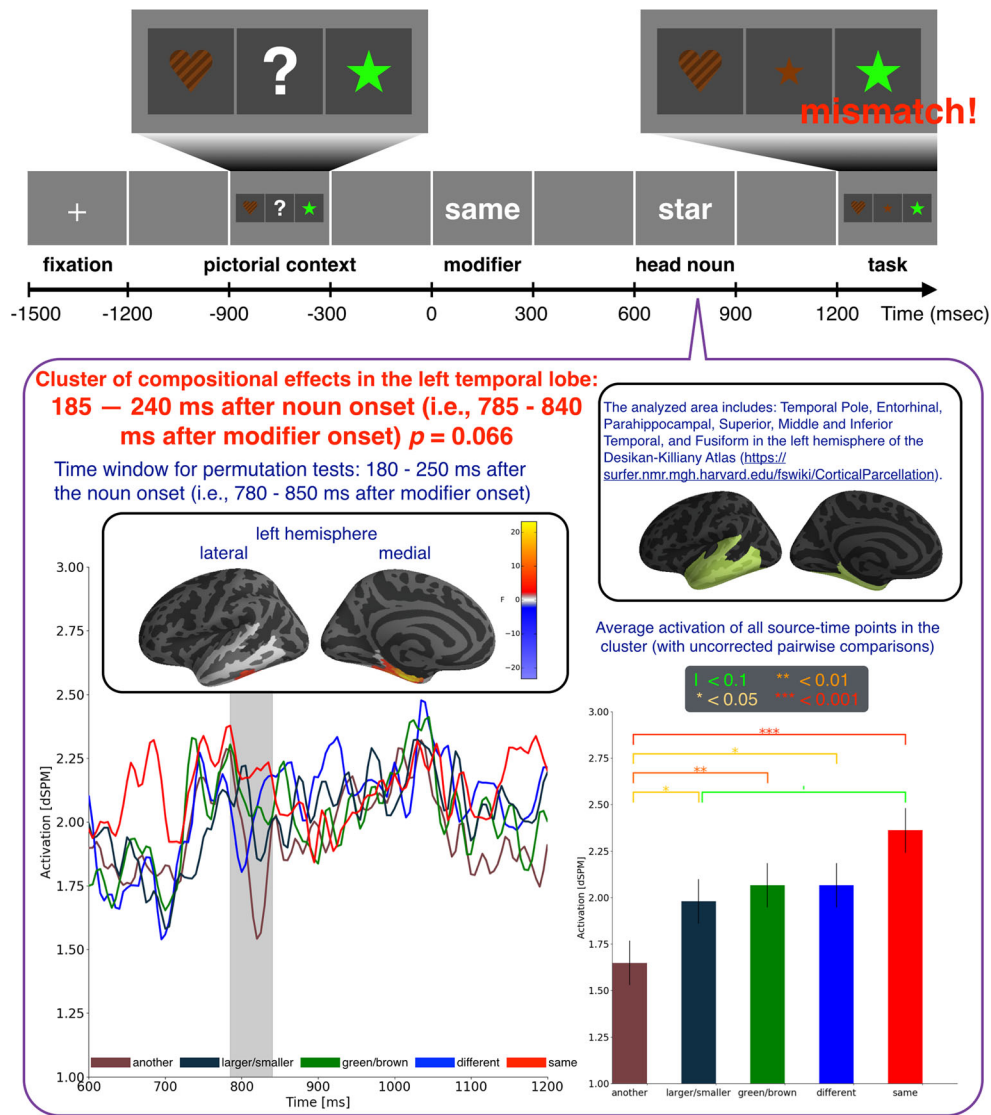


Fig. 1 Trial structure and results of clustering tests for compositional effects. The top row in the figure shows the trial structure. Under that is the marginally significant cluster of compositional effects in the left

temporal lobe during head noun presentation: from 185 to 240 ms after noun onset ($p = .066$). This main effect was driven by increased activity for the *same*, *different*, color, and comparative NP conditions.

of modifiers and nouns, and four filler trials. The trials within each block were randomized and the blocks were also randomized for each participant. Between adjacent blocks, participants could choose to take a short rest or continue immediately.

Procedure

Prior to each MEG recording, we used a Polhemus FastSCAN three-dimensional laser digitizer to scan the participant’s head shape and locate the positions of five marker coils placed across the forehead. The digitized head shape was later used to constrain source localization during data processing by co-registering the position of the coils with respect to the MEG sensors.

Before the MEG recording, there was a practice session (containing 17 trials) outside the magnetically shielded room. During this practice, participants were given the same instructions as in the MEG recording, and for each trial they got feedback after their responses, so that they could verify their comprehension of the task.

During the MEG recording, participants lay in a dimly lit magnetically shielded room. The positions of the marker coils were measured at the beginning and the end of the experiment. MEG data were collected by using a whole-head 208-channel axial gradiometer system (Kanazawa Institute of Technology, Nonoichi, Japan), at a 1000-Hz sampling rate with a low-pass filter at 200 Hz. Stimuli were projected onto a screen about 50 cm away from participants’ eyes. During the recording,

participants got no feedback after responses. The recording session lasted approximately 65 min.

Data processing

The raw MEG data were first noise-reduced via the continuously adjusted least-squares method (Adachi, Shimogawara, Higuchi, Haruta, & Ochiai, 2001) in the MEG Laboratory software 2.004A (Yokogawa Electric and Eagle Technology Corp., Japan), before being processed and analyzed with MNE-Python (Gramfort et al., 2013; Gramfort et al., 2014) and Eelbrain package 0.22.1 (<http://pythonhosted.org/eelbrain/>).

The function of independent component analysis (with the method of “fastica,” the parameter of accounting for 95% of the raw data, and a rejection level of 3,000 fT/cm) was applied in order to remove known artifacts (e.g., heartbeats, eye blinks) (<http://martinos.org/mne/dev/manual/preprocessing/ica.html>). Afterward, the MEG data were low-pass filtered at 40 Hz, epoched (i.e., segmented out) from 100 ms before the onset of the pictorial context to 600 ms after the head noun. Individual epochs were automatically rejected if any sensor value exceeded 2,000 fT/cm at any time point. If the raw activity for individual channels in an epoch greatly deviated from neighboring channels, activity for those channels in that epoch was manually removed, and field interpolation was applied. Epochs were baseline-corrected with the prestimulus interval: from –100 ms to the onset of the pictorial context.

We co-registered each participant’s head shape with the standard FreeSurfer average brain—fsaverage (CorTech and MGH/HMS/MIT Athinoula A. Martinos Center for Biomedical Imaging)—in order to construct their cortical surface. Cortically constrained minimum norm estimates (Hämäläinen & Ilmoniemi, 1994) were then calculated with MNE with a standard procedure. After constructing each participant’s cortical surface, the boundary element method was adopted to calculate the forward solution (Mosher, Leahy, & Lewis, 1999). Then the inverse solution was computed from the forward solution together with the noise covariance matrix (from the 100-ms prestimulus interval). The inverse solution was applied to the evoked data, yielding current estimates at each source for three orthogonal dipoles, which represented the most likely distribution of neural activity. We chose to retain the length of the resulting current vector—that is, the orientation of the current dipoles was not fixed with relation to the cortical surface—and thus the brain activation is reported without reference to orientation. The resulting minimum norm estimates of neural activity were noise-normalized at each spatial location to reduce the location bias of the estimates (Dale et al., 2000), and the resulting time- and location-dependent values yielded dynamic statistical parameter maps (dSPM), providing information about the statistical reliability of the estimated signal at each location with millisecond accuracy. dSPM values were the input for

spatiotemporal clustering tests. Source data were down-sampled to one observation every 5 ms in order to achieve computational tractability in our spatiotemporal permutation clustering tests.

Statistical analyses

Behavioral data We checked the overall accuracy for all 528 trials to verify whether participants paid enough attention. Then the behavioral data from the 480 critical trials were analyzed in a one-way repeated measures analysis of variance (ANOVA) for accuracy and speed. Reaction times were measured from the onset of the postverbal picture for each correctly answered trial and for each participant. Reaction times above three *SDs* from each participant’s own average were discarded from the reaction time analysis.

Spatiotemporal clustering tests A whole-brain comparison of the MEG data between the consonant and color NPs revealed that, as we had anticipated, this consonant string did not work as a baseline. The consonant NP activated a large area in the left hemisphere, much more than the color NP did, suggesting that in the present paradigm, the consonant string had some type of disruptive effect on processing. Thus, the consonant NP was discarded as a baseline.

On the basis of spatiotemporal information from previous MEG studies on LATL compositional effects for two-word NPs (Bemis & Pykkänen, 2011; Zhang & Pykkänen, 2015; Ziegler & Pykkänen, 2016), we performed a cluster-based one-way permutation ANOVA across our five combinatorial conditions within the time window 180–250 ms after the noun onset (which slightly extended the previously detected peaking time 200–250 ms in order to fully reveal the temporal extent of potential clusters) in the left temporal lobe centering around the LATL (including temporal pole and the entorhinal, parahippocampal, fusiform, and superior, middle, and inferior temporal gyri, as in the Desikan–Killiany Atlas; see <https://surfer.nmr.mgh.harvard.edu/fswiki/CorticalParcellation> and Desikan et al., 2006), testing for spatiotemporal clusters of brain activation that showed statistically significant differences among our conditions, corrected for multiple comparisons (Maris & Oostenveld, 2007). Since there were more trials for the conditions with comparative and color NPs, an equalized number of trials for all conditions were randomly selected out using Eelbrain’s parameter “equalize-evoked-counts.”

Our spatiotemporal permutation clustering test was performed in Eelbrain 0.22.1 with a standard procedure. For each source at each time interval (i.e., every 5 ms), an uncorrected ANOVA was calculated. If a significant effect at a *p* value of .05 (uncorrected) was observed in at least ten contiguous sources for at least 25 ms, these data points were treated as a cluster, and the *F* values of each data point within the cluster

were summed as a test statistic. By randomly reassigning the conditional labels of the dSPM data within participants 10,000 times, we obtained a set of test statistics to form the null distribution. Each cluster of interest was assigned a corrected p value ($\alpha = .05$) based on this null distribution.

Similarly, to explore semantics-related effects in the AG/SMG, we used a combination of the supramarginal and inferior parietal gyri (see Binder & Desai, 2011; Demonet et al., 1992) in the left hemisphere of the Desikan–Killiany Atlas to perform spatiotemporal clustering tests. Since previous studies offered no temporal information for reference, to maximize our ability to detect possible effects for future studies, we divided the whole timeline into 150-ms time windows (i.e., 0–150 ms, 150–300 ms, etc.) and ran clustering tests. When a detected cluster occurred at the boundary, we extended the time window for 50 ms and reperformed the test. The effects reported are not corrected for multiple comparisons across these time windows and should be treated as preliminary.

Finally, to test retrieval/dependency-related effects in the LIFG, we used the combination of pars opercularis and pars triangularis (e.g., Cooke et al., 2002) in the left hemisphere of the Desikan–Killiany Atlas and the time window of 300–600 ms after the head noun onset (Leiken et al., 2015; Leiken & Pylkkänen, 2014) to perform a spatiotemporal clustering test.

Whole-brain pairwise comparison Since our spatiotemporal clustering test yielded a marginally significant cluster in the time window of 185–240 ms after the noun onset, to visualize how widespread the effects were for each type of modifier, we further conducted pairwise whole-brain comparisons between all of the main conditions of interest, specifically: (1) comparative NP versus *another* NP, (2) color NP versus *another* NP, (3) *different* NP versus *another* NP, (4), *same* NP versus *another* NP, (5) *same* NP versus comparative NP, (6) *same* NP versus color NP, and (7) *same* NP versus *different* NP. The results were plotted when at least ten adjacent sources showed a difference at a p value of .05 (uncorrected for multiple comparisons) for at least 25 ms.

Similarly, for the clusters yielded from the spatiotemporal clustering tests investigating the AG/SMG and the LIFG, we also conducted pairwise whole-brain comparisons between all of the main conditions of interest during the time window of the clusters.

Results

Behavioral data

The mean accuracy over the 480 critical trials for the 21 participants whose data were included in the MEG analyses was 98.01% ($SD = 1.44\%$). The one-way repeated

measures ANOVA revealed a significant difference among trials containing different modifiers [$F(7, 160) = 5.534, p < .0001$]. A Tukey's post-hoc test revealed that the accuracy for trials with *different* was significantly lower than the accuracy for trials with any other modifier, even though the mean accuracy for *different* NP was also as high as 95%. The one-way repeated measures ANOVA on the reaction time data revealed no significant difference among trials containing different modifiers [$F(7, 160) = 1.505, p = .17$]. Overall, the performance was very good, and the reaction times were similar. Thus, we consider that the task was overall equally demanding across all the conditions. The mean accuracies and reaction times for each modifier are summarized in Table 2.

Compositional effects: Spatiotemporal clustering results and whole-brain pairwise comparisons

Figure 1 shows the marginally significant cluster (185–240 ms after the noun onset, $p = .066$) elicited in our spatiotemporal cluster-based one-way ANOVA.

Planned pairwise comparisons in this cluster showed that this main effect of condition was driven by increased activity for the *same*, *different*, color, and comparative NP conditions. Specifically, they all elicited higher amplitudes than the *another* NP (*same* NP, $p < .001$; *different* NP, $p = .02$; color NP, $p = .004$; comparative NP, $p = .03$); the *same* NP condition was also increased relative to the comparative NP ($p = .071$), *different* NP ($p = .103$), and color NP ($p = .104$) with a marginal significance. Overall, *same*, *different*, and even gradable adjectives in their comparative form all engaged combinatory activity in the ventral/medial part of the left temporal lobe as much as did the color adjectives, and *same* engaged this brain region the most.

Figure 2 depicts the whole-brain pairwise comparisons, confirming that *same*, *different*, and color NPs engaged the LATL more than did *another* NPs; however, although the effects touched the left temporal pole, they were mainly distributed over more medial left temporal cortex, including the entorhinal cortex, parahippocampal gyrus, fusiform gyrus, and part of inferior temporal cortex. The comparison between comparative and *another* NPs shows that the effects elicited by comparative NPs were rather in the posterior part of the inferior temporal lobe. The comparison between *same* and *different* NPs shows that *same* NPs engaged the middle to posterior parts of the temporal lobe more than did *different* NPs.

Pairwise comparisons also showed that *same*, *different*, and color NPs engaged the left occipital lobe more than did *another* NPs, but in these plots, the clusters shown in the left occipital lobe are not contiguous with those in the left temporal lobe.

Table 2 Behavioral data: Means and standard deviations of accuracy (as percentages) and reaction times (in seconds) for the 480 critical trials

Condition	Example	Means and SDs of Accuracy (%)	Means and SDs of Reaction Times (s)
Same NP	<i>same star</i>	98.02% (2.39%)	0.7971 (0.2221)
Different NP	<i>different star</i>	95.00% (4.86%)	0.9627 (0.3361)
Color NP	<i>green star</i>	98.25% (1.62%)	0.7607 (0.2421)
	<i>brown star</i>	98.89% (1.52%)	0.7832 (0.2457)
Comparative NP	<i>larger star</i>	98.81% (1.41%)	0.7671 (0.2281)
	<i>smaller star</i>	97.86% (2.42%)	0.8762 (0.3159)
Another NP	<i>another star</i>	98.49% (1.97%)	0.8187 (0.2997)
Consonant NP	<i>ykdmjz star</i>	98.73% (1.74%)	0.7603 (0.1936)

Overall, these whole-brain comparisons suggest that the marginally significant cluster observed in the spatiotemporal permutation ANOVA is indeed focal to the searched brain area and not part of another, larger cluster.

Relationality effects: Spatiotemporal clustering results and whole-brain pairwise comparisons

Within the searched time window of 450–650 ms after modifier onset, we found a significant cluster at 515–620 ms ($p = .038$) (see Fig. 3). Within this cluster, planned pairwise

comparisons showed that this main effect was driven by increased activity for comparative NPs. Specifically, comparative NPs elicited higher amplitudes than did color NPs ($p = .018$), different NPs ($p < .001$), and same NPs ($p = .017$), as well as marginally higher amplitudes than another NPs ($p = .091$). The amplitudes for different NPs were also lower than those for another NPs ($p = .017$), color NPs ($p = .015$), and same NPs ($p = .007$). Whole-brain pairwise comparisons between the conditions of interest showed that these elicited effects were located mainly in the left SMG, spreading toward the left AG (see Fig. 4).

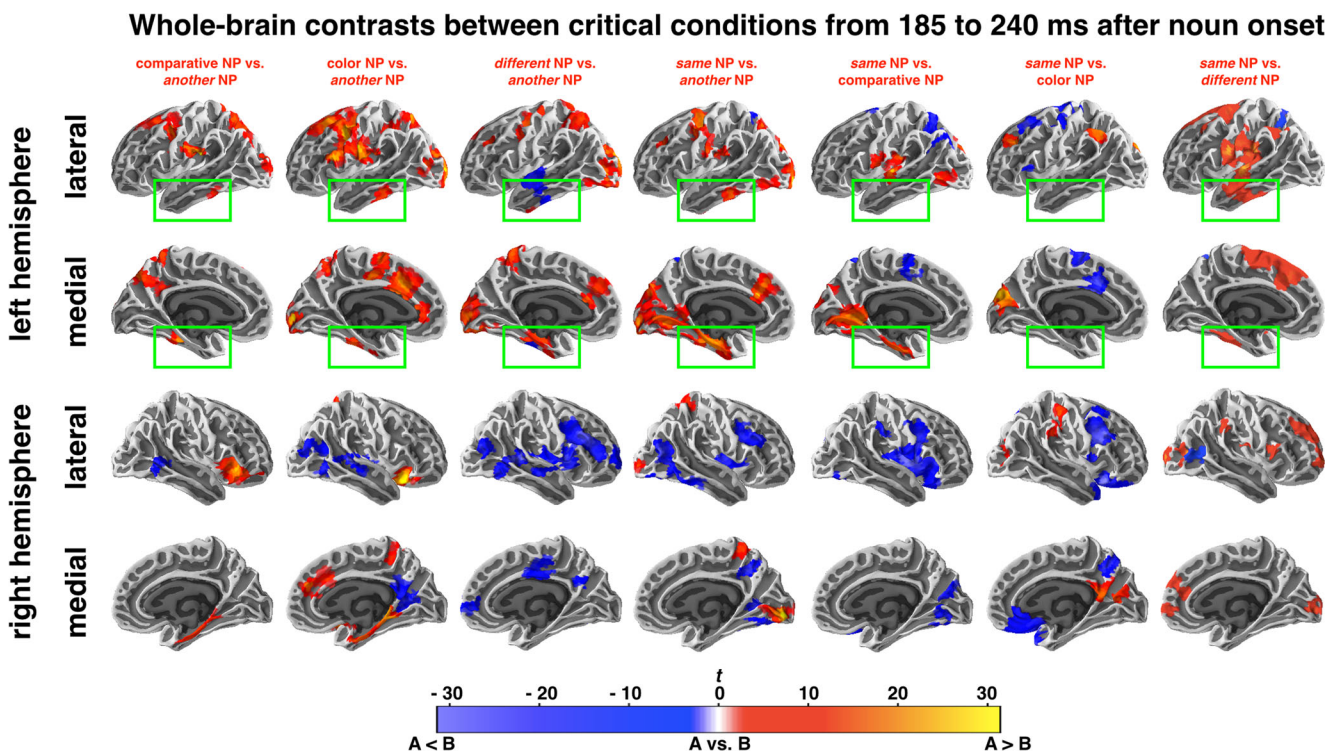


Fig. 2 Whole-brain contrasts between the critical conditions from 185 to 240 ms after noun onset. From left to right: (1) comparative NP versus another NP, (2) color NP versus another NP, (3) different NP versus another NP, (4) same NP versus another NP, (5) same NP versus comparative NP, (6) same NP versus color NP, and (7) same NP versus

different NP. Each plot shows all sources that were part of a cluster at any time during 185 and 240 ms after noun onset. Each cluster is formed on the basis of at least ten data points showing a significant effect at a p value of .05 (uncorrected) for at least 25 ms.

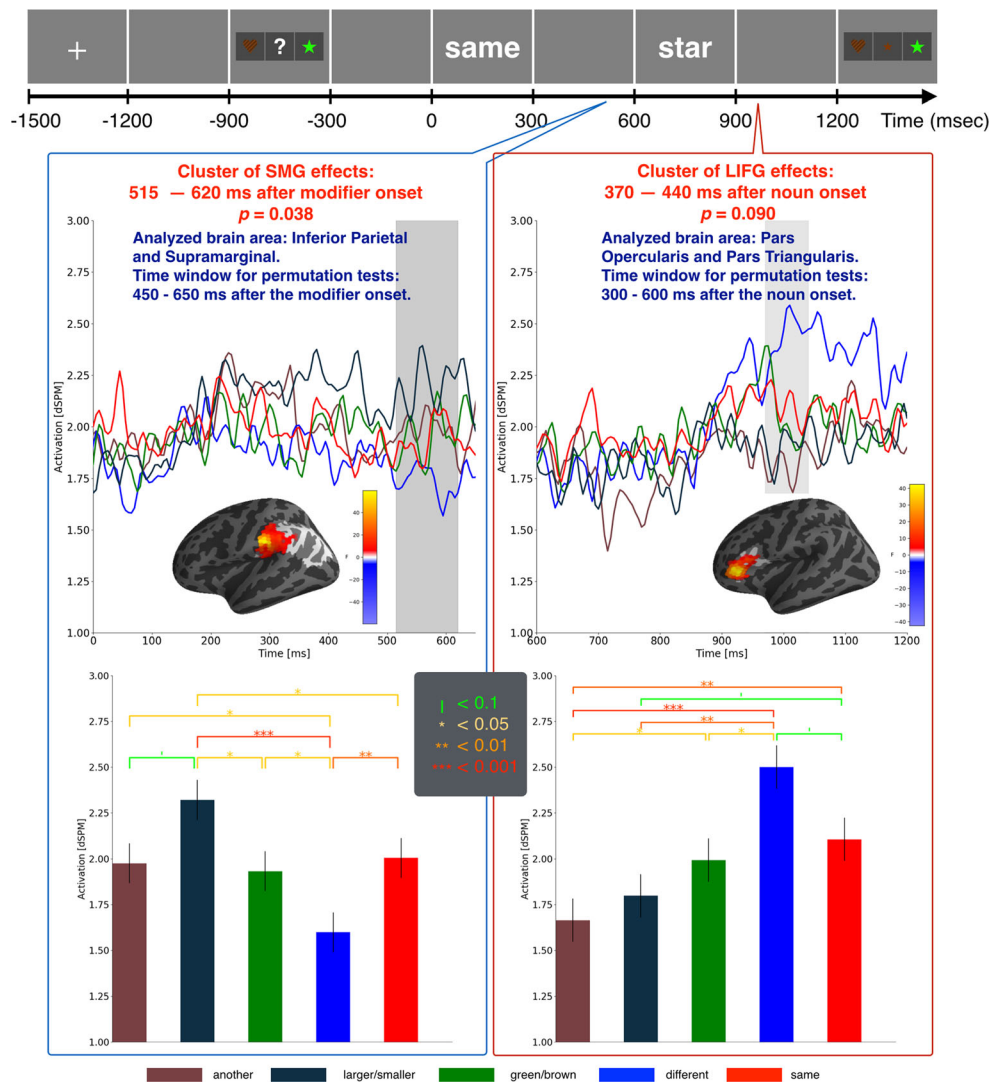


Fig. 3 Results of clustering tests for supramarginal gyrus (SMG) and left inferior frontal gyrus (LIFG) effects. The left panel shows the significant cluster of SMG effects during modifier presentation: from 515 to 620 ms ($p = .038$). For this cluster, the main effect was driven by increased SMG activity for comparative NPs. The right panel shows a cluster with a trend toward LIFG effects during noun presentation: from 370 to 440 ms ($p = .09$). For this cluster, the main effect was most driven by increased LIFG

activity for *different* NPs. The bar graphs indicate the average activations of all source time points in their corresponding cluster (with uncorrected pairwise comparisons). The whitened parts in the brains indicate the analyzed brain areas: the combination of inferior parietal gyrus and SMG in the left hemisphere of the Desikan–Killiany Atlas, for SMG effects (left), and the combination of pars opercularis and pars triangularis in the left hemisphere of the Desikan–Killiany Atlas for LIFG effects (right).

Retrieval effects: Spatiotemporal clustering results and whole-brain pairwise comparisons

The spatiotemporal clustering test yielded a cluster for a trend toward LIFG effects during noun presentation: from 370 to 440 ms ($p = .09$; see Fig. 3). Planned pairwise comparisons showed that the effect within this cluster was driven mainly by increased activity for *different* NPs, which elicited larger amplitudes than did *another* NPs ($p < .001$), comparative NPs ($p = .002$), color NPs ($p = .029$), and *same* NPs ($p = .06$). In addition, *same* NPs also elicited larger amplitudes than did *another* NPs ($p = .003$) and comparative NPs ($p = .081$), and color NPs elicited larger amplitudes than did *another* NPs ($p =$

.018). Whole-brain pairwise comparisons between the conditions of interest showed that these elicited effects were indeed mainly located in the LIFG, spreading toward a more anterior part in some comparisons (see Fig. 5).

Discussion

In this study we investigated the neural mechanism underlying the processing of *same* and *different*. By adding a pictorial context before the canonical two-word NP paradigm, we were able to investigate the processing of context-dependent NPs along with context-independent NPs. Our main result was that

Whole-brain contrasts between critical conditions from 515 to 620 ms after modifier onset

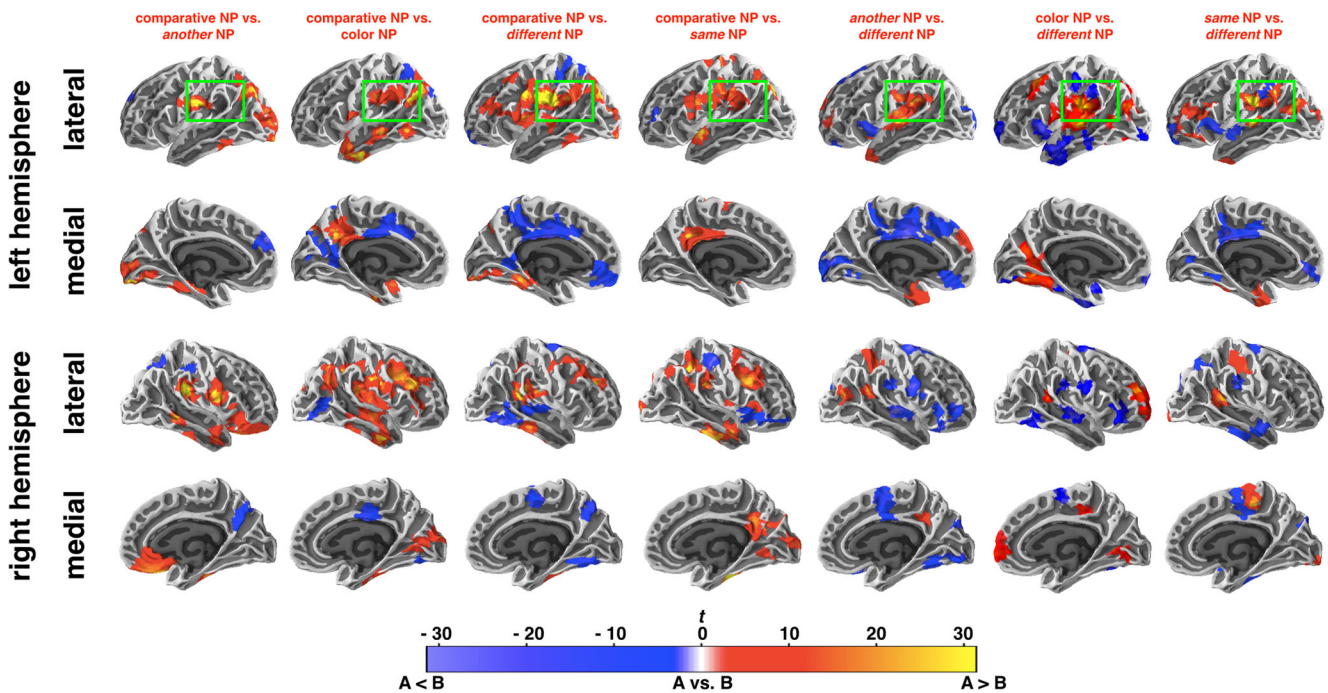


Fig. 4 Whole-brain contrasts between the critical conditions from 515 to 620 ms after modifier onset. From left to right: (1) comparative NP versus *another* NP, (2) comparative NP versus color NP, (3) comparative NP versus *different* NP, (4), comparative NP versus *same* NP, (5) *another* NP versus *different* NP, (6) color NP versus *different* NP, and (7) *same* NP

versus *different* NP. Each plot shows all sources that were part of a cluster at any time from 515 to 620 ms after modifier onset. Each cluster is formed on the basis of at least ten data points showing a significant effect at a p value of .05 (uncorrected) for at least 25 ms.

Whole-brain contrasts between critical conditions from 370 to 440 ms after noun onset

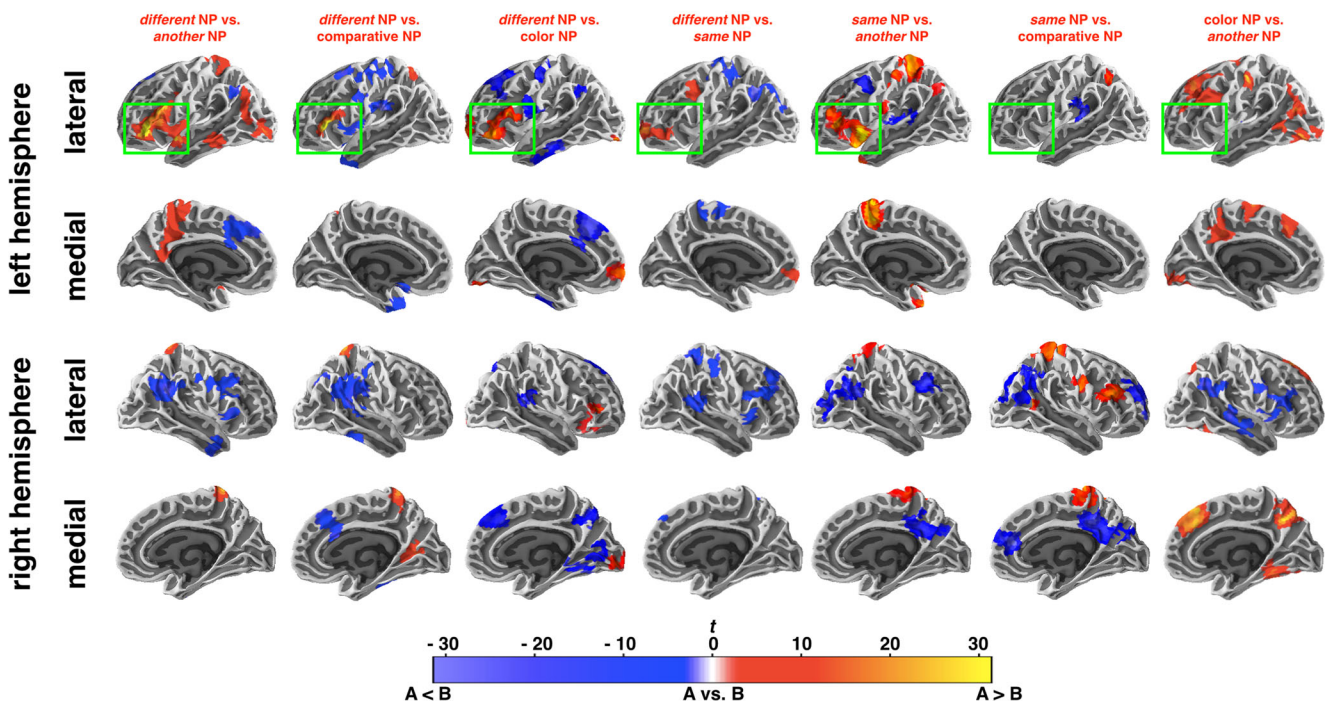


Fig. 5 Whole-brain contrasts between the critical conditions from 370 to 440 ms after noun onset. From left to right: (1) *different* NP versus *another* NP, (2) *different* NP versus comparative NP, (3) *different* NP versus color NP, (4), *different* NP versus *same* NP, (5) *same* NP versus *another* NP, (6) *same* NP versus comparative NP, and (7) color NP versus

another NP. Each plot shows all sources that were part of a cluster at any time from 370 to 440 ms after noun onset. Each cluster is formed on the basis of at least ten data points showing a significant effect at a p value of .05 (uncorrected) for at least 25 ms.

although *same* and *different* lack featural information in their word meanings, they elicited compositional effects in the left temporal lobe as early as lexical adjectives did, and the effects for *same* were even larger than those for the colors.

Although the temporal details of the present results are similar to those of previous MEG studies investigating LATL compositional effects in the processing of two-word NPs, the localization is more posterior and medial. Overall, our cluster implicated a brain region similar to the results of Bemis and Pykkänen (2013a, 2013b), Westerlund et al. (2015), Westerlund and Pykkänen (2014), and Ziegler and Pykkänen (2016), but more posterior and medial than the brain region implicated in the results of Bemis and Pykkänen (2011) and Zhang and Pykkänen (2015). One clear difference between the present study and prior work was the limited number of stimuli in each condition, a necessary design feature given our aim to study specific functional vocabulary items, which do not partake in large classes. Further, unlike in prior work, in which pictures or words have been used as comprehension prompts after the verbal stimuli, our word stimuli were preceded by pictures. This also could have engaged slightly distinct cortical circuits due to top-down context effects. But what mattered the most for our hypotheses was whether the functional adjectives patterned similarly to or differently from the lexical ones, and our results clearly showed similarity.

In terms of time windows, our cluster was also slightly earlier (i.e., 185–240 after the noun onset) than those elicited in previous studies (roughly 200–250 ms after the noun onset), which might have been due to priming effects that resulted from our repetitive use of modifiers and nouns. In addition to this similarity in spatiotemporal details, within this cluster, all conditions involving featural information (i.e., *same*, *different*, color, and comparative NPs) elicited more activity than the baseline *another* NPs. Thus, the cluster observed in the present study is likely to reflect LATL compositional effects.

It is worth noting that in previous studies, size-related gradable adjectives did not always elicit LATL compositional effects as color terms did, and contextual information with regard to the comparison class of the noun seemed necessary for size-related gradable adjectives to elicit LATL effects (see Ziegler & Pykkänen, 2016, vs. Poortman & Pykkänen, 2016). Our results, together with those of Poortman and Pykkänen (2016), seem to confirm that when a pictorial context is presented before linguistic stimuli, so that the comparison class of the noun can be inferred and size-related gradable adjectives interpreted in a less vague way, these gradable adjectives can elicit compositional effects as color terms do.

Therefore, consistent with prior work, we showed combinatory effects in the left temporal lobe and have offered the first example of composition with a functional word as one of the input words. Thus, although compositional effects in the

left temporal lobe—as far as we know today—do not reflect aspects of composition that do not add features, such as quantification (Blanco-Elorrieta, 2016; Del Prato & Pykkänen, 2014), they can be elicited by functional words that carry no featural information themselves, but reference conceptual features in context.

In our supplementary analyses of the left AG and SMG, the result patterns within our elicited cluster suggested that the processing of comparatives indeed elicits increased activity in the left SMG, whereas the processing of *same* and *different* did not activate these regions as much. Presumably, *same* and *different* address the homogeneity or heterogeneity among featural information, whereas comparatives relate entities in a certain dimension. The cognitive details underlying the use of these words, the spatiotemporal details of relevant effects in the SMG in semantic processing, as well as the cognitive functions of these brain regions still require future investigations.

As for the LIFG, the results of our supplementary analyses regarding retrieval effects conform in timing with prior LIFG results for the MEG correlates of syntactic dependency formation (Leiken et al., 2015; Leiken & Pykkänen, 2014). The increased activation in the LIFG during the processing of *different* and *same* NPs suggests that the semantic retrieval of featural information from contexts also involves the LIFG. Put together, these findings support a general view of the LIFG as a retrieval/dependency-related brain region.

Intriguingly, *different* NPs activated LIFG the most. Presumably, *different* can be considered the negation of *same*, and thus the processing of *different* might involve more computation and be more working memory demanding. This possibility is compatible with the hypothesis that the LIFG is a working-memory-relevant brain region (Barde, Schwartz, & Thompson-Schill, 2006; Cohen et al., 1997; Paulesu, Frith, & Frackowiak, 1993; Rogalsky et al., 2008). More specifically, in the present design, when interpreted in a context in which there was, say, a middle-sized green striped star, the expression *different star* rules out the possibility that the star under discussion is of the same color, size, and pattern. Thus, with the use of this hidden negation in the meaning of *different*, the increase in LIFG activity elicited by the processing of *different* may also reflect a selection of information among competing alternatives (see Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997).

Another intriguing issue is that our LIFG retrieval effects were temporally later than compositional effects. One possibility is that the retrieval of specific featural information and the formation of a holistic mental representation might constitute a dynamic process, and by the time that compositional effects occur in the LATL, *same* (i.e., the word that drove compositional effects most in the present study) might have just provided information that the features on multiple relevant dimensions needed to be taken into consideration (see

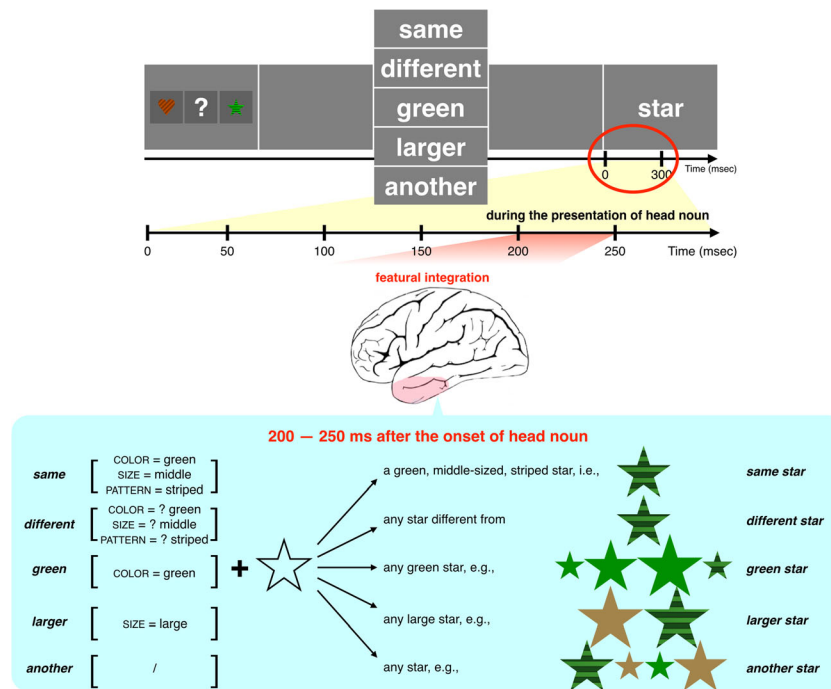


Fig. 6 Featural integration in the left temporal lobe. In processing a combinatorial noun phrase of the form “modifier + head noun,” about 200–250 ms after onset of the head noun, following word recognition and the processing of modality-specific attributes stored in a widely

distributed neural network, featural information is integrated by the left temporal lobe in order to form a holistic mental representation for the semantics of the noun phrase.

also Fig. 6). Another possibility is that, as was proposed by Thompson-Schill et al. (1997), the later LIFG effects reflect the selection, not retrieval, of semantic information, and thus the increased activity for *different* (i.e., the word that drove LIFG activity most in the present study) reflects a selection process that follows semantic composition in the left temporal lobe; that is, at this point, the composed representation rules out a set of possibilities. These two possibilities are, however, not incompatible with each other, and future work will be needed to flesh out the details.

To summarize, the present study contributes to our existing understanding of the function of the left temporal lobe in language processing and explores the potential division of labor in different brain areas and different time windows. The left anterior temporal lobe (especially its ventral part; see Murphy et al., 2017) has been hypothesized to be a semantic “hub,” integrating modality-specific attributes stored in a widely distributed neural network into an amodal holistic mental representation (Patterson et al., 2007). Figure 6 illustrates how the processing of *same* and *different* can be viewed as parallel to the processing of lexical adjectives within this semantic “hub” model. *Same* transmits features on all relevant dimensions from a given context, and thus can bring highly specific featural information in an efficient way, whereas *different* is more vague and conveys some nondeterminism—one or more of the context-provided values are subject to changes, so that a specific kind of possibility is ruled out. Eventually, featural

information is composed with the meaning of the head noun, leading to a compositional semantic representation.

Overall, our findings suggest that overlapping neural mechanisms are involved in the processing of both lexical adjectives that carry features themselves and functional adjectives that recruit and transmit featural information from a context. Since our context was nonlinguistic, these results suggest that the mechanism underlying semantic processing in the brain may compose information from a variety of sources—that is, from our words and from the world in which we live.

Supplementary Materials

The scripts and materials of this study are available on the website of NYU Neurolinguistics Lab: https://psych.nyu.edu/pylkkannen/lab/linminzhang_supplementarymaterials.html.

Author note This research was funded by the National Science Foundation, Grant BCS-1221723 (to L.P.), and by Grant G1001 from the NYUAD Institute, New York University Abu Dhabi (also to L.P.). The authors thank Graham Flick for data collection. The authors declare no competing financial interests.

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*, 669–672. doi:<https://doi.org/10.1109/77.919433>
- Ardila, A., Concha, M., & Rosselli, M. (2000). Angular gyrus syndrome revisited: Acalculia, finger agnosia, right–left disorientation and semantic aphasia. *Aphasiology*, *14*, 743–754.
- Barbieri, E., Aggujaro, S., Molteni, F., & Luzzatti, C. (2015). Does argument structure complexity affect reading? A case study of an Italian agrammatic patient with deep dyslexia. *Applied Psycholinguistics*, *36*, 533–558. doi:<https://doi.org/10.1017/S0142716413000337>
- Barde, L. H. F., Schwartz, M. F., & Thompson-Schill, S. (2006). The role of left inferior frontal gyrus (LIFG) in semantic short-term memory: A comparison of two case studies. *Brain and Language*, *99*, 82–83.
- Bemis, D. K., & Pyllkkänen, L. (2011). Simple composition: A magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *Journal of Neuroscience*, *31*, 2801–2814. doi:<https://doi.org/10.1523/JNEUROSCI.5003-10.2011>
- Bemis, D. K., & Pyllkkänen, L. (2013a). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex*, *23*, 1859–1973.
- Bemis, D. K., & Pyllkkänen, L. (2013b). Flexible composition: MEG evidence for the deployment of basic combinatorial linguistic mechanisms in response to task demands. *PLoS ONE*, *8*, e73949. doi:<https://doi.org/10.1371/journal.pone.0073949>
- Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., & Grodzinsky, Y. (2003). The neural reality of syntactic transformations: Evidence from functional magnetic resonance imaging. *Psychological Science*, *14*, 433–440.
- Ben-Shachar, M., Palti, D., & Grodzinsky, Y. (2004). Neural correlates of syntactic movement: Converging evidence from two fMRI experiments. *NeuroImage*, *21*, 1320–1336.
- Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in Cognitive Sciences*, *15*, 527–536. doi:<https://doi.org/10.1016/j.tics.2011.10.001>
- Binney, J. R., Embleton, K. V., Jefferies, E., Parker, G. J. M., & Lambon Ralph, M. A. (2010). The ventral and inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: Evidence from a novel direct comparison of distortion-corrected fMRI, rTMS, and semantic dementia. *Cerebral Cortex*, *20*, 2728–2738.
- Blanco-Elorrieta, E., & Pyllkkänen, L. (2016). Composition of complex numbers: Delineating the computational role of the left anterior temporal lobe. *NeuroImage*, *124*, 194–203.
- Bonner, M. F., Peelle, J. E., Cook, P. A., & Grossman, M. (2013). Heteromodal conceptual processing in the angular gyrus. *NeuroImage*, *71*, 175–186. doi:<https://doi.org/10.1016/j.neuroimage.2013.01.006>
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436. doi:<https://doi.org/10.1163/156856897X00357>
- Brennan, J., & Pyllkkänen, L. (2012). The time-course and spatial distribution of brain activity associated with sentence processing. *NeuroImage*, *60*, 1139–1148.
- Caplan, D., Alpert, N., Waters, G., & Olivieri, A. (2000). Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping*, *9*, 65–71.
- Caplan, D., Stanczak, L., & Waters, G. (2008). Syntactic and thematic constraint effects on blood oxygenation level dependent signal correlates of comprehension of relative clauses. *Journal of Cognitive Neuroscience*, *20*, 643–656. doi:<https://doi.org/10.1162/jocn.2008.20044>
- Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, *386*, 604–608. doi:<https://doi.org/10.1038/386604a0>
- Constable, R. T., Pugh, K. R., Berroya, E., Mencl, W. E., Westerveld, M., Ni, W., & Shankweiler, D. (2004). Sentence complexity and input modality effects in sentence comprehension: An fMRI study. *NeuroImage*, *22*, 11–21.
- Cooke, A., Zurif, E. B., DeVita, C., Alsop, D., Koenig, P., Detre, J., . . . Grossman, M. (2002). Neural basis for sentence comprehension: Grammatical and short-term memory components. *Human Brain Mapping*, *15*, 80–94.
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., & Halgren, E. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, *26*, 55–67.
- Damasio, H., Grabowski, T. J., Tranel, D., Ponto, L. L., Hichwa, R. D., & Damasio, A. R. (2001). Neural correlates of naming actions and of naming spatial relations. *NeuroImage*, *13*, 1053–1064.
- Del Prado, P., & Pyllkkänen, 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. doi:<https://doi.org/10.3389/fpsyg.2014.00524>
- Demonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., . . . Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, *115*, 1753–1768.
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., . . . Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*, 968–980.
- de Zubicaray, G. I., Hansen, S., & McMahon, K. L. (2013). Differential processing of thematic and categorical conceptual relations in spoken word production. *Journal of Experimental Psychology: General*, *142*, 131–142. doi:<https://doi.org/10.1037/a0028717>
- Dronkers, N. F., Wiilkins, D. P., Vanvalin, R. D., Redfern, B. B., & Jaeger, J. J. (1994). A reconsideration of the brain-areas involved in the disruption of morphosyntactic comprehension. *Brain and Language*, *47*, 461–463.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., . . . Hämäläinen, M. S. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 267.
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., . . . Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, *86*, 446–460.
- Graves, W. W., Binder, J. R., Desai, R. H., Conant, L. L., & Seidenberg, M. S. (2010). Neural correlates of implicit and explicit combinatorial semantic processing. *NeuroImage*, *53*, 638–646.
- Hämäläinen, M. S., & Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: Minimum norm estimates. *Medical and Biological Engineering and Computing*, *32*, 35–42.
- Hart, J., & Gordon, B. (1990). Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Annals of Neurology*, *27*, 226–231.
- Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, *115*, 1783–1806.
- Humphries, C., Binder, J. R., Medler, D., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience*, *18*, 665–679.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2007). Time course of semantic processes during sentence comprehension: An fMRI study. *NeuroImage*, *36*, 924–932.
- 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*, 128–138.

- Humphries, C., Willard, K., Buchsbaum, B., & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: An fMRI study. *NeuroReport*, *12*, 1749–1752.
- Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thurlborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science*, *274*, 114–116.
- Keller, T. A., Carpenter, P. A., & Just, M. A. (2001). The neural bases of sentence comprehension: A fMRI examination of syntactic and lexical processing. *Cerebral Cortex*, *11*, 223–237.
- Lau, E., Almeida, D., Hines, P. C., & Poeppel, D. (2009). A lexical basis for N400 context effects: Evidence from MEG. *Brain and Language*, *111*, 161–172. doi:<https://doi.org/10.1016/j.bandl.2009.08.007>
- Leiken, K., McElree, B., & Pykkänen, L. (2015). Filling predictable and unpredictable gaps, with and without similarity-based interference: Evidence for LIFG effects of dependency processing. *Frontiers in Psychology*, *6*, 1739. doi:<https://doi.org/10.3389/fpsyg.2015.01739>
- Leiken, L., & Pykkänen, L. (2014). MEG evidence that the LIFG effect of object extraction requires similarity-based interference. *Language, Cognition and Neuroscience*, *29*, 381–389.
- Lewis, G. A., Poeppel, D., & Murphy, G. L. (2015). The neural bases of taxonomic and thematic conceptual relations: An MEG study. *Neuropsychologia*, *68*, 176–189. doi:<https://doi.org/10.1016/j.neuropsychologia.2015.01.011>
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190.
- Martinho, A., & Kacelnik, A. (2016). Ducklings imprint on the relational concept of “same or different.” *Science*, *353*, 286–288.
- Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., . . . Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*, 467–479.
- Mosher, J. C., Leahy, R. M., & Lewis, P. S. (1999). EEG and MEG: Forward solutions for inverse methods. *IEEE Transactions on Biomedical Engineering*, *46*, 245–259.
- Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S. J., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, *47*, 36–45.
- Mummery, C. J., Patterson, K., Wise, R. J. S., Vandenberg, R., Price, C. J., & Hodges, J. R. (1999a). Disrupted temporal lobe connections in semantic dementia. *Brain*, *122*, 61–73.
- Mummery, C. J., Shallice, T., & Price, C. J. (1999b). Dual-process model in semantic priming: A functional imaging perspective. *NeuroImage*, *9*, 516–525.
- Murphy, C., Rueschemeyer, S.-A., Watson, D., Karapanagiotidis, T., Smallwood, J., & Jefferies, E. (2017). Fractionating the anterior temporal lobe: MVPA reveals differential responses to input and conceptual modality. *NeuroImage*, *147*, 19–31.
- Newman, A. J., Pancheva, R., Ozawa, K., Neville, H. J., & Ullman, M. T. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of Psycholinguistic Research*, *30*, 339–364.
- Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (2013). Going beyond inferior prefrontal involvement in semantic control: Evidence for the additional contribution of dorsal angular gyrus and posterior middle temporal cortex. *Journal of Cognitive Neuroscience*, *25*, 1824–1850.
- Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences*, *108*, 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*, 976–987. doi:<https://doi.org/10.1038/nrn2277>
- Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature*, *362*, 342–345.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442. doi:<https://doi.org/10.1163/156856897X00366>
- Poortman, E. B., & Pykkänen, L. (2016). Adjective conjunction as a window into the LATL’s contribution to conceptual combination. *Brain and Language*, *160*, 50–60. doi:<https://doi.org/10.1016/j.bandl.2016.07.006>
- Price, A. R., Bonner, M. F., Peelle, J. E., & Grossman, M. (2015). Converging evidence for the neuroanatomic basis of combinatorial semantics in the angular gyrus. *Journal of Neuroscience*, *35*, 3276–3284.
- Price, C. J. (2010). The anatomy of language: A review of 100 fMRI studies published in 2009. *Annals of the New York Academy of Sciences*, *1191*, 62–88.
- Rogalsky, C., Matchin, W., & Hickok, G. (2008). Broca’s area, sentence comprehension, and working memory: An fMRI study. *Frontiers in Human Neuroscience*, *2*, 14. doi:<https://doi.org/10.3389/fpsyg.2008.014.2008>
- Seghier, M. L. (2013). The angular gyrus multiple functions and multiple subdivisions. *The Neuroscientist*, *19*, 43–61.
- Seghier, M. L., Fagan, E., & Price, C. J. (2011). Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. *Journal of Neuroscience*, *30*, 16809–16817.
- Stowe, L. A., Broere, C. A. J., Paans, A. M. J., Wijers, A. A., Mulder, G., Vaalburg, W., & Zwarts, F. (1998). Localizing components of a complex task: Sentence processing and working memory. *NeuroReport*, *9*, 2995–2999.
- Stromswold, K., Caplan, D., Alpert, N., & Rauch, S. (1996). Localization of syntactic comprehension by positron emission tomography. *Brain and Language*, *52*, 452–473.
- Thompson, C. K., Bonakdarpour, B., Fix, S. C., Blumenfeld, H. K., Parish, T. B., Gitelman, D. R., & Mesulam, M.-M. (2007). Neural correlates of verb argument structure processing. *Journal of Cognitive Neuroscience*, *19*, 1753–1767. doi:<https://doi.org/10.1162/jocn.2007.19.11.1753>
- Thompson-Schill, S. L., D’Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences*, *94*, 14792–14797.
- Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience*, *14*, 550–560.
- Westerlund, M., Kastner, I., Al Kaabi, M., & Pykkänen, L. (2015). The LATL as locus of composition: MEG evidence from English and Arabic. *Brain and Language*, *141*, 124–134.
- Westerlund, M., & Pykkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia*, *57*, 59–70.
- Williams, A., Reddigari, S., & Pykkänen, L. (2017). Early sensitivity of left perisylvian cortex to relationality in nouns and verbs. *Neuropsychologia*, *100*, 131–143. doi:<https://doi.org/10.1016/j.neuropsychologia.2017.04.029>
- Zhang, L. (2016). External and internal *same*: A unified account motivated by attitude reports. *Sinn und Bedeutung*, *20*, 833–850.
- Zhang, L., & Pykkänen, L. (2015). The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. *NeuroImage*, *111*, 228–240. doi:<https://doi.org/10.1016/j.neuroimage.2015.02.028>
- Ziegler, J., & Pykkänen, L. (2016). Scalar adjectives and the temporal unfolding of semantic composition: An MEG investigation. *Neuropsychologia*, *89*, 161–171.