Neuropsychologia 89 (2016) 161–171

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Scalar adjectives and the temporal unfolding of semantic composition: An MEG investigation

Jayden Ziegler[a,*], Liina Pylkkänen[a,b,c]

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

ARTICLE INFO

Article history:
Received 31 March 2016
Received in revised form
19 May 2016
Accepted 9 June 2016
Available online 11 June 2016

Keywords:
Scalar adjectives
Context-sensitivity
Semantic composition
Left anterior temporal lobe (LATL)
Magnetoencephalography (MEG)

ABSTRACT

A growing body of research implicates the left anterior temporal lobe (LATL) for combinatorial semantic processing. However, magnetoencephalography (MEG) studies have revealed this activity to be timed quite early, at 200–250 ms, preceding the most common time window for lexical-semantic effects. What type of semantic composition could the LATL perform at 200–250 ms? We hypothesized that the LATL computes an early stage of composition, taking as its input only the most readily available lexical-semantic information. To test this, we varied the context-sensitivity of prenominal adjectives, postulating that only context-insensitive intersective adjectives (e.g., *dead*, *Italian*) should compose in an early time window, whereas the composition of context-sensitive scalar adjectives (e.g., *fast*, *large*) should be delayed until the interpretation of the subsequent noun is fully determined. Consistent with this, early combinatory effects in left temporal cortex were observed only for intersective adjectives, though in this study the effects were somewhat more posterior than in prior reports. Overall, our results suggest multiple stages of semantic composition, of which the LATL may index the earliest.

1. Introduction

Language allows us to communicate infinite complex ideas by means of a finite inventory of words. In the past two decades, cognitive neuroscience research has begun to characterize the processes underlying this expressive power of language, consistently implicating the left anterior temporal lobe (LATL) as a major locus of sentence-level combinatory linguistic operations (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Humphries et al., 2005, 2006; Xu et al., 2005; Humphries et al., 2007; Rogalsky and Hickock, 2009; Baron et al., 2010; Brennan et al., 2010; Baron and Osherson, 2011; Pallier et al., 2011; Brennan and Pylkkänen, 2012). More recently, a series of magnetoencephalography (MEG) studies has further constrained this interpretation, implicating the LATL in very basic phrase building specifically (Bemis and Pylkkänen, 2011, 2012, 2013; Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014; Westerlund and Pylkkänen, 2014; Westerlund et al., 2015).

Despite such advancements, a key question that remains is the precise processing level of LATL-localizing combinatory activity. Timing-wise, the basic composition effects in MEG (i.e., increased amplitude for combinatory two-word phrases vs. non-combinatory controls) have consistently emerged for adjective-noun pairs between 200 and 250 ms post-noun onset (Bemis and Pylkkänen, 2011, 2012; Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014; Westerlund and Pylkkänen, 2014; Westerlund et al., 2015). Moreover, despite their somewhat early occurrence—namely, just after the visual M170 (or "Type II" activity), a response known to show a preference for letter strings (Tarkiainen et al., 1999) but no sensitivity to semantic variables, at least in isolated words (Simon et al., 2012)—recent evidence also suggests that these computations may in fact be semantic, as opposed to syntactic, in nature. Specifically, within syntactically parallel expressions, the combinatory effect in the LATL has shown sensitivity to several semantic factors, including the conceptual specificity of the composing items (Westerlund and Pylkkänen, 2014), the degree to which the composition results in a complex concept as opposed to a complex meaning more generally (Del Prato and Pylkkänen, 2014), and the relevance of the composition to reference resolution (Leffel et al., 2014).

Since the meanings of lexical items cannot be combined before they have been accessed, these findings are intriguing, especially in light of the fact that although some evidence exists for very early single-word lexical-semantic effects (behavioral: Marslen-Wilson, 1973; Marslen-Wilson and Tyler, 1975; Marslen-Wilson,

* Correspondence to: Dept. of Psychology, Harvard University, 33 Kirkland St., Cambridge, MA 02138, United States.
E-mail addresses: ziegler@g.harvard.edu (J. Ziegler), liina.pylkkanen@nyu.edu (L. Pylkkänen).

http://dx.doi.org/10.1016/j.neuropsychologia.2016.06.010
0028-3932/

1985, 1987; Rastle et al., 2000; Mohr and Pulvermüller, 2002; electrophysiological: Pulvermüller et al., 2001; Shtyrov and Pulvermüller, 2002; Marinkovic et al., 2003; Endrass et al., 2004; Hinojosa et al., 2004; Shtyrov et al., 2004; Pulvermüller, 2005; Pulvermüller et al., 2005; Shtyrov et al., 2005; Shimotake et al., 2015; Chen et al., 2016), the most typical time window for semantic priming or other lexical-semantic effects is somewhat later, at 300–400 ms post-stimulus onset (Fischler et al., 1983; Smith and Halgren, 1987; Rugg, 1990; Holcomb and McPherson, 1994; Federmeier and Kutas, 1999; Pylkkänen et al., 2002; Laszlo and Federmeier, 2010; for reviews, see Kutas and Van Petten, 1988, 1994; Kutas and Federmeier, 2000; Pylkkänen and Marantz, 2003; Van Petten and Luka, 2006; Federmeier, 2007; Lau et al., 2008, 2009; Kutas and Federmeier, 2009, 2011). Thus, a question arises: What could be the functional contribution of LATL combinatory activity such that it precedes the bulk of lexical-semantic effects?

One possibility, consistent with the rather long time window associated with lexical-semantic effects, is that instead of occurring at a single processing stage, semantic activation is in fact gradient (e.g., Moss, 1997), perhaps unfolding over hundreds of milliseconds. This general idea resonates with an extant proposal by Binder and Desai (2011) on the neurobiological architecture of semantic memory. Within their "embodied abstraction" framework, conceptual access progresses from highly schematic representations to more detailed ones, and not all levels are necessarily activated under all task demands. Correspondingly, one could imagine combinatory operations at various stages of this gradient activation, taking as input representations that have been specified to varying degrees over the course of semantic access and processing (Pylkkänen, 2015).

The main goal of the present work was to provide an initial test of this type of hypothesis. Specifically, we aimed to vary the level of detail at which the currently processed word (i.e., the noun) needs to be interpreted in order to compose with its modifier (i.e., the adjective), the logic being that, at the LATL stage around 200 ms, we should only see effects of combinatory operations that do not require a high level of semantic detail. To achieve this, we employed a well-studied variable in the formal semantics literature on adjectives—namely, the context-sensitivity of scalar adjectives.

The meanings of context-sensitive adjectives, such as *fast* and *large*, depend heavily on the nouns they combine with (Kamp, 1975; Klein, 1980; Bierwisch, 1989; Kennedy, 1999; Kennedy and McNally, 2005; Kennedy, 2007, 2012). For example, *large* does not mean much unless we know what the relevant "comparison class" is—in other words, "large for what"? Thus, a *large elephant* and a *large mouse* are radically different in size, although in both cases the same word *large* is used to describe them. In contrast, an intersective adjective, such as *dead* or *Italian*, is relatively context-insensitive; that is, it has a more constant meaning across uses, and consequently, its meaning is relatively specific even in isolation (Kamp, 1975; Kamp and Partee, 1995; Partee, 1995; Kennedy, 1999, 2012). Thus, in prenominal position, context-insensitive adjectives receive an interpretation at the adjective, independent of the following noun, while context-sensitive adjectives *require* the following noun for interpretation. As such, for context-sensitive adjectives only, an additional step of meaning computation is required—namely, at the noun, the adjective's meaning must first be computed before the meaning of the phrase can be composed. For context-insensitive adjectives, on the other hand, the phrasal meaning is composed immediately at the noun. For present purposes, we were interested, specifically, in whether the LATL would exhibit distinct temporal profiles for adjectives with varying degrees of inherent context-sensitivity.

To operationalize our intended contrast between intersective and scalar adjectives, we employed the classic *for*-phrase test, yielding well-formed expressions for scalars (e.g., *large/tall for a desk*) and ill-formed ones for intersectives (**wooden/*Italian for a desk*; Kamp, 1975; Siegel, 1976; Kamp and Partee, 1995). It is important to note that this test taps directly into the *scalarity* of the adjective, as opposed to context-sensitivity in a more general sense. For example, although *Italian for a car* and *Italian for a vacation* sound ill-formed, which for the present purposes categorizes *Italian* as intersective, the interpretation of *Italian* is clearly somewhat different in *Italian car* (i.e., most likely a car made in Italy) and *Italian vacation* (i.e., a vacation taking place in Italy). In fact, although intersectives and scalars are typically taken to contrast in context-sensitivity, few, if any, adjectives are entirely context-insensitive. The crucial assumption for the present purposes is that upon encountering *Italian*, the comprehender should at least be able to commit to the reading *somehow relating to Italy*, whereas for scalars such as *large*, the interpretation at the adjective is less specified. While ultimately an empirical question, our current aim was to test whether early combinatory activity in the LATL would be lessened or eliminated for scalar as compared to intersective adjectives, following the, perhaps controversial, assumption that scalars are more context-sensitive than intersectives.

As a secondary aim of this study, we also sought to provide initial evidence for the time window at which context-sensitive scalar adjectives might compose, should they indeed fail to show evidence of early composition in the LATL. As a potentially relevant variable, we employed the conceptual specificity of the post-adjectival nouns, following Westerlund and Pylkkänen (2014). The conceptual specificity manipulation was intended as a way to affect the size of the comparison class provided for the adjective, more specific nouns providing a narrower class. We reasoned that the time point at which the specificity of the noun affects the processing of the scalar phrase is conceivably a time at which the noun and adjective meanings are combined. Thus, we varied noun specificity to assess whether scalar modification is sensitive to the size of the comparison class at some later time.

One remaining challenge in this literature concerns the anatomical variability in the localization of various semantic and/or combinatory effects across left anterior temporal cortex. In general, both the temporal pole (Brodmann area 38; Gauthier et al., 1997; Grabowski et al., 2001; Bright et al., 2004; Pobric et al., 2007; Baron and Osherson, 2011; Clarke et al., 2013; Del Prato and Pylkkänen, 2014) and middle and ventral temporal gyri (Brodmann areas 20 and 21; Gauthier et al., 1997; Grabowski et al., 2001; Bright et al., 2004; Tyler et al., 2004; Rogers et al., 2006; Pobric et al., 2007; Baron and Osherson, 2011; Clarke et al., 2011, 2013; Westerlund and Pylkkänen, 2014) have been implicated, but the computational distinctions between these regions remain elusive. In the present work, we did not aim to address this outstanding issue but rather employed Brodmann area (BA) 21 as our region of interest (ROI), given that due to its central location within left anterior temporal cortex and the somewhat blurry spatial resolution of MEG, it is likely to capture activity from each of the potentially relevant regions.

To assess the anteriority of our effects within BA 21, which also covers posterior temporal cortex, we complemented our ROI analysis with full brain contrasts visualizing activity source by source. Crucially, the specific aim of our study was to address the earliness of LATL-localizing combinatory activity in particular, and thus, we did not analyze any ROIs not covering left temporal cortex. In other words, this study was not conceived of as a general exploration of the effects of context-sensitivity or semantic composition in the brain, but rather as a targeted investigation of the effect of the scalar vs. intersective contrast on early LATL activity. Thus, eventually, the present findings will need to be incorporated into a broader understanding of the semantic network in general, which is known to encompass many additional regions, such as the angular gyrus (e.g., Bonner et al., 2013), left inferior frontal cortex (e.g., Rodd et al., 2005), and ventromedial prefrontal cortex (e.g., Bemis and Pylkkänen, 2011), among others. Here, we did not, however, aim to explore this network more widely, but rather focused specifically on explaining the early timing of the LATL combinatory response, well characterized in prior studies. Consequently, our analyses will not rule out any accounts of other semantic regions.

In sum, our study varied both the context-sensitivity of adjectives to test whether highly context-sensitive adjectives would fail to show early combinatory effects in the LATL, and the conceptual specificity of the subsequent nouns as a possible way to obtain evidence for a later combinatory stage for such context-sensitive adjectives.

2. Methods

2.1. Participants

Twenty-seven right-handed, native English speakers participated in our experiment. All participants were non-colorblind with normal or corrected-to-normal vision and provided written consent prior to participating. Three participants were excluded due to low accuracy in the behavioral task ($<80\%$), leaving 24 participants in the final analysis (17 female; mean age=24, SD=4).

2.2. Materials

The present experiment was a 2×3 design with Adjective Type (Scalar, Intersective, and No Adjective) and Noun Type (LowSpec and HighSpec) as factors (Table 1). For the nouns, we generated 50 noun pairs, one of each pair representing a more general, or low-specificity, category (LowSpec; e.g., *dish*), and the other representing a more specific, or high-specificity, example thereof (HighSpec; e.g., *bowl*). In all cases, the high-specificity noun was in a set-theoretic subset relation to the low-specificity noun (e.g., the set of *bowls* is a proper subset of the set of *dishes*), selected on the basis of hypernym/hyponym categorization in WordNet (Fellbaum, 1998). LowSpec and HighSpec nouns were matched for length (LowSpec mean=6.12, SD=1.91; HighSpec mean=5.66, SD=1.85; $t[98]=1.22$, $p=.22$), number of morphemes (NMorph; LowSpec mean=1.24, SD=.43; HighSpec mean=1.20, SD=.40; $t[98]=.48$, $p=.63$), and lexical decision reaction time (LDRT; LowSpec mean=632.27, SD=63.66; HighSpec mean=642.72, SD=68.19; $t[98]=-.79$, $p=.43$; values from the English Lexicon Project, Balota et al., 2007). Nouns were not matched for log HAL frequency (LowSpec mean=9.31, SD=1.46; HighSpec mean=8.67, SD=1.28; $t[98]=2.32$, $p=.02$; values from the English Lexicon Project), however, as doing so would have significantly constrained the number of possible superset-subset pairs included in the experiment, though the frequency difference was not large (Table 2).

Each of the target words was paired, in turn, with a scalar adjective (ScalLowSpec, e.g., *large dish*; ScalHighSpec, e.g., *large bowl*), an intersective adjective (IntLowSpec, e.g., *wooden dish*; IntHighSpec, e.g., *wooden bowl*), or an unpronounceable consonant string (LowSpec[OneWord], e.g., *srbgfn dish*; HighSpec[OneWord], e.g., *xcvhwf bowl*), resulting in a total of 300 unique stimuli. Following Pylkkänen and colleagues (Bemis and Pylkkänen, 2011, 2012; Westerlund and Pylkkänen, 2014; Westerlund et al., 2015), the use of unpronounceable consonant strings in the No Adjective condition was to ensure that the amount of pre-noun visual stimulation across all conditions was as closely matched as possible. To this end, for any given noun, consonant strings were matched in character length to the mean of the corresponding scalar and intersective adjectives in the two-word conditions.

For the adjectives, we initially generated separate lists of scalars and intersectives based both on the theoretical literature and intuition, which were then later submitted to a norming study on Amazon Mechanical Turk (www.mturk.com) for verification. The norming task consisted of asking 88 participants (51 female; mean age=31, SD=10) to judge the well-formedness of our two adjective types in combination with our two noun types in simple *for*-phrases (i.e., "[adj.] for a [noun]" constructions), in which scalar adjectives are perfectly licit while intersective adjectives are not (e.g., *large for a dish*, **wooden for a bowl*; Kamp, 1975; Siegel, 1976; Kamp and Partee, 1995). Participants indicated their responses using a 7-point Likert scale (1=completely unnatural, 7=perfectly natural). This task was approved by New York University's Institutional Review Board, and informed consent was obtained from all participants prior to their participation.

In our analysis of these results, we computed a linear mixed-effects model on judgment ratings in R using the lmer function in the lme4 package (Bates, 2010). Fixed effects included Adjective Type, Noun Type, their interaction, Noun Frequency, and Transitional Probability (from the adjective to the noun; see below), and participant and item were treated as random intercepts. We found a significant main effect of Adjective Type, $t(132)=13.44$, $p<.001$, such that participants, on average, judged our scalar adjectives in these constructions to be more natural (mean=5.44, SD=1.74) than our intersective adjectives (mean=2.54, SD=1.81). No other main effects or interactions were found.

Like the nouns, the adjectives were not matched on log HAL frequency (Scalar mean=10.47, SD=1.55; Intersective mean=8.72, SD=1.59; $t[48]=3.91$, $p<.001$; values from the English Lexicon Project), as doing so would have severely limited the number and breadth of items included in the stimulus set.

Final stimuli were matched for transitional probability from the adjective to the noun (ScalLowSpec mean=.0011, SD=.0031; ScalHighSpec mean=.0005, SD=.0009; IntLowSpec mean=.0027, SD=.0142; IntHighSpec mean=.0003, SD=.0010; Adjective Type main effect: $F[1,196]=.62$, $p=.57$; Noun Type main effect: $F[1,196]=2.82$, $p=.34$; interaction: $F[1,196]=.76$, $p=.38$; values calculated from the Google Books [American English] Corpus, Davies, 2011) and for Latent Semantic Analysis (LSA) distance between the adjective and the noun (ScalLowSpec mean=.21, SD=.15; ScalHighSpec mean=.23, SD=.13; IntLowSpec mean=.13, SD=.15; IntHighSpec mean=.11, SD=.13; Adjective Type main effect: $F[1,196]=35.85$, $p=.11$; Noun Type main effect: $F[1,196]=.03$,

Table 1
Experimental design.

		Noun Type	
		LowSpec	HighSpec
Adj. Type	Scalar	*large dish*	*large bowl*
	Intersective	*wooden dish*	*wooden bowl*
	No Adjective	*srbgfn dish*	*xcvhwf bowl*

Table 2
Summary of noun statistics (values from the English Lexicon Project).

	No.	Length		Freq.		NMorph		LDRT	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
LowSpec	50	6.12	1.91	9.31	1.46	1.24	.43	632.27	63.66
HighSpec	50	5.66	1.85	8.67	1.28	1.20	.40	642.72	68.19
t-test *p*		.22		**.02**		.63		.43	

Table 3
Summary of bigram statistics (values from the Google Books [American English] Corpus and Latent Semantic Analysis @ CU Boulder).

	No.	Trans. Prob.		LSA	
		Mean	SD	Mean	SD
ScalLowSpec	50	.0011	.0031	.21	.15
ScalHighSpec	50	.0005	.0009	.23	.13
IntLowSpec	50	.0027	.0142	.13	.15
IntHighSpec	50	.0003	.0010	.11	.13
ANOVA *p* (Adj. Type)		.57		.11	
ANOVA *p* (Noun Type)		.34		.88	
ANOVA *p* (Interaction)		.38		.41	

p=.88; interaction: F[1,196]=.68, p=.41; values from Latent Semantic Analysis @ CU Boulder, Landauer et al., 1998; Table 3). We chose to focus on transitional probabilities over simple bigram frequencies for two reasons. First, as a linguistic measure, transitional probability includes bigram frequency in its calculation (Miller and Selfridge, 1950), and the two are thus correlated. Second, transitional probabilities have been shown to be a powerful tool in the acquisition of both words (Aslin et al., 1998) and syntax (Thompson and Newport, 2007), above and beyond simple bigram co-occurrence, making them an important factor in guiding linguistic parsing more generally.

2.3. Procedure

Prior to the MEG recordings, we used a Polhemus Fastscan three-dimensional laser digitizer (Polhemus, Inc., Vermont, USA) to determine the shape of participants' heads, as well as the orientation of three marker coils across the forehead and two on the tragi of the ears. These measures were later used to constrain source localization of the elicited activity during data processing by orienting the position of the participant's head with respect to the MEG sensors. Participants completed a practice block of items prior to beginning the experiment.

For the actual experiment, participants lay in a dimly lit magnetically-shielded room while stimuli were projected onto a screen approximately 50 cm away from their eyes. MEG data were collected using a 157-channel whole-brain axial gradiometer system (Kanazawa Institute of Technology, Nonoichi, Japan) at a sampling rate of 1000 Hz with a 200 Hz low-pass filter and 60 Hz notch filter. The positions of the marker coils were measured at both the beginning and end of the experiment.

The experiment consisted of 300 trials spread across 10 blocks (i.e., 30 trials/block). Blocks were constructed such that no two adjectives or nouns in a given block were the same. Both the stimuli within the blocks and the blocks themselves were presented in random order using PsychToolBox (Brainard, 1997; Kleiner et al., 2007). Each word appeared onscreen for 300 ms in 30-point white Courier font on a gray background. A blank screen was presented for 300 ms between words, as well as between the final word and following task. Task questions consisted of standalone words or phrases (e.g., *platter*) that had to be judged as either matching or mismatching the preceding adjective-noun phrase (e.g., *large dish*), and were chosen to engage semantic processing of the combined phrases specifically (i.e., they could not generally be answered on the basis of either the adjective or noun in isolation). These remained onscreen until participants responded by pressing a button with either the index or middle finger of their left hand (Fig. 1). No feedback was provided during the experiment. Within a given experimental block, exactly half of the questions were correct, and the other half were incorrect. Task questions for individual stimuli remained constant across participants, and all participants saw all stimuli. Each session, including preparation, practice, and recording, lasted approximately 45 minutes.

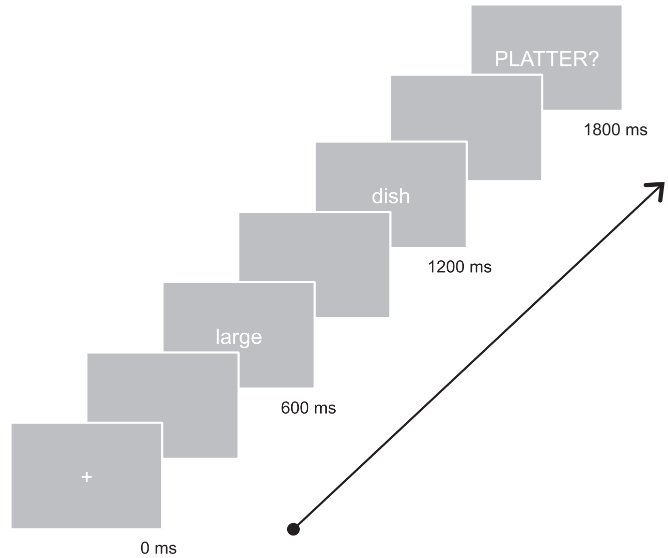

Fig. 1. Trial structure.

2.4. Data analysis

2.4.1. Noun time window

MEG data were collected continuously and subsequently epoched for each trial from 700 ms pre-noun onset to 600 ms post-noun onset. This long interval captured activity elicited both at the adjective and at the noun. A 1 Hz high-pass filter was applied prior to epoching. We then removed artifacts in the data by rejecting trials in which the maximum amplitude during our epoch of interest exceeded 3000 fT, or when participants blinked. Eye blinks were determined by visual inspection of each trial. These steps resulted in a loss of 24.54% of trials overall.

Raw data were then averaged for each condition and low-pass filtered at 40 Hz. Source activity was estimated using separate distributed L2 minimum norm estimates (Hämäläinen and Ilmoniemi, 1994) for each averaged condition for each participant using BESA 5.1 (MEGIS Software, GmbH, Gräfelfing, Germany). Source estimates were computed by placing two shells, each containing 713 evenly distributed regional sources, at 10–30% below a smoothed standard brain provided by BESA. The 713 BESA sources were all assigned Brodmann area-level labels automatically using the Talairach Daemon (1988 Talairach Atlas; Lancaster et al., 2000) on the basis of their source coordinates in Talairach space. For each trial, activity was time-locked to noun onset, and we defined the activity baseline (i.e., the channel noise covariance matrix used for computing the minimum norm estimates) as the 100 ms window pre-noun onset.

To analyze the time course of LATL activity, we aimed to select a single Brodmann area (BA) most likely to capture the relevant combinatory effects, as implicated by prior studies. Though the precise localization of LATL composition effects has varied somewhat along the temporal pole and more lateral left anterior cortex, the best convergence of relevant results can be identified in the middle temporal gyrus (BA 21). Indeed, BA 21 has consistently been implicated as a major locus of combinatorial semantic processing with materials similar to ours (Westerlund and Pylkkänen, 2014; Westerlund et al., 2015), as well as being implicated in much of the fMRI work showing differences in LATL sensitivity for more vs. less specific concepts, in both healthy participants (Gauthier et al., 1997; Grabowski et al., 2001; Bright et al., 2004; Tyler et al., 2004; Rogers et al., 2006; Pobric et al., 2007; Clarke et al., 2011; Clarke et al., 2013) and those with LATL atrophy (Schwartz et al., 1979; Snowden et al., 1989; Hodges et al., 1992; Hodges et al., 1995; Mummery et al., 1999, 2000; Garrard and Hodges, 2000; Gorno-Tempini et al., 2004; Rogers et al., 2004, 2006; Gainotti, 2006; Garrard and Carroll, 2006; Patterson et al., 2006; Gainotti, 2007, 2012), which is particularly relevant to our noun specificity contrast. Finally, recent work bridging these two literatures (Westerlund and Pylkkänen, 2014) found its strongest effects at roughly the anterior portion of left BA 21. Thus, based on these literatures, we chose left BA 21 as our region of interest (ROI). Crucially, however, we additionally used follow-up uncorrected full brain contrasts (see below) to visualize the precise spatial distribution of our effects, given that BA 21 of course extends all the way to posterior temporal cortex as well, and thus would not reflect only anterior activity. Our prediction was that our BA 21 effects should primarily be driven by activity in the anterior portion of BA 21.

ROI source data were first subjected to a nonparametric, cluster-based permutation test (Maris and Oostenveld, 2007) aimed at identifying temporal clusters of activity reflecting our experimental manipulation, corrected for multiple comparisons. As the main test statistic, we employed a 2×2 repeated-measures ANOVA on the combinatory conditions alone, directly assessing contrasts between the two adjective types (Scalar vs. Intersective) and how they are affected by noun specificity (HighSpec vs. LowSpec). For cluster selection, samples showing an effect at an uncorrected level of $p < .3$ were grouped into clusters when ten or more adjacent samples showed such an effect, as in Bemis and Pylkkänen (2011). The cluster with the largest summed test statistic was then isolated, and from 10,000 random permutations, a corrected p-value ($\alpha < .05$) was generated as the ratio of permutations yielding a higher test statistic than the actual observed test statistic.

The permutation test was conducted over a window of activity spanning 200–500 ms post-stimulus onset in order to capture both early combinatory activity, shown in previous studies to peak at roughly 250 ms post-noun onset (Bemis and Pylkkänen, 2011, 2012; Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014; Westerlund and Pylkkänen, 2014; Westerlund et al., 2015), and possible later combinatory activity associated with the processing of scalar adjectives in particular, as potentially reflected by an effect of the size of the comparison class provided for the adjective (i.e., noun specificity).

Reliable main effects and interactions were then unpacked with sample-by-sample pairwise cluster-based permutation t-tests, also corrected for multiple comparisons, using the same time interval and cluster selection thresholds as for the ANOVA (see above).

Finally, targeted one-tailed permutation t-tests were conducted to identify early (200–300 ms) LATL composition effects for each two-word condition as compared to its one-word control (i.e., IntLowSpec vs. LowSpec[OneWord], ScalLowSpec vs. LowSpec[OneWord], IntHighSpec vs. HighSpec[OneWord], ScalHighSpec vs. HighSpec[OneWord]). Since this analysis only tested for previously reported early effects, our search for clusters was limited to the 200–300 ms interval.

Lastly, given that our noun manipulation was motivated by prior literature implicating increased LATL amplitudes for higher-specificity single nouns, we examined the effect of specificity within the non-combinatory one-word conditions (i.e., HighSpec [OneWord] vs. LowSpec[OneWord]) in a one-tailed permutation t-test over the entire 200–500 ms time window of interest.

As already mentioned above, ROI analyses were followed by uncorrected pairwise t-tests over the full brain to verify that our observed ROI effects reflected activity within BA 21 rather than adjacent regions, and to examine the spatial distributions of our effects within BA 21 more specifically. In this analysis, spatio-temporal clusters were required to maintain significance ($p < .05$) across at least five adjacent sources and for at least five consecutive milliseconds.

2.4.2. Adjective time window

We performed a post hoc analysis at adjective onset comparing scalar to intersective adjectives to determine whether the effects observed in our main analyses may instead have been driven by an inherent difference between the two adjective types individually. The activity baseline was now defined as the 100 ms window pre-adjective onset. 71.78% of the overall trials were included in the analysis. For this analysis, we performed a two-tailed permutation t-test across the entire adjective time window, from adjective onset to noun onset (i.e., -600 to 0 ms).

3. Results

3.1. Behavioral results

Mean accuracy across all 24 participants on all 300 conditions was 89.54% (SD=3.70%), and the mean response time was 1.61 s (SD=.62 s). As the task in the present experiment was intended solely to ensure attention and not specifically to tap into the computations under study, no further analyses were conducted on the behavioral results.

3.2. Noun time window

3.2.1. ROI results

The 2×2 ANOVA on the combinatory conditions revealed an early main effect of Adjective Type (200–317 ms; $p = .05$; Fig. 2(a)), such that phrases with intersective adjectives exhibited a higher average amplitude than those with scalar adjectives (Fig. 2(b)), consistent with our predictions. The increase for intersectives did not interact with Noun Type in the ANOVA, and consistent with this, planned pairwise permutation t-tests confirmed a similar, though non-significant, early increase for Intersective over Scalar adjectives within the HighSpec (i.e., IntHighSpec vs. ScalHighSpec) and LowSpec nouns (i.e., IntLowSpec vs. ScalLowSpec) individually (Fig. 2(c)).

The main analysis also revealed a later significant interaction between Adjective Type and Noun Type (350–471 ms; $p = .04$; Fig. 2(a)), with condition means suggesting an effect of noun specificity within the scalar but not within the intersective conditions (Fig. 2(b)). This was confirmed by the pairwise permutation t-tests, which revealed a reliable cluster of increased activity at 333–460 ms for the HighSpec over LowSpec nouns when modified by scalars (i.e., ScalHighSpec vs. ScalLowSpec; $p = .02$), but no such cluster for the same comparison within the intersectives (i.e., IntHighSpec vs. IntLowSpec; Fig. 2(c)).

Finally, early composition effects were assessed for all combinations of Adjective Type and Noun Type by comparing each two-word condition to its one-word control, revealing exactly the pattern that would be predicted on the basis of Westerlund and Pylkkänen's (2014) prior findings (i.e., early composition effects for LowSpec nouns only). Specifically, we observed an increase for two-word over one-word conditions for LowSpec nouns with intersective modifiers at 200–258 ms (IntLowSpec vs. LowSpec[OneWord]; $p = .047$), while no such clusters were found for any of the other pairwise comparisons (i.e., ScalLowSpec vs. LowSpec[OneWord], IntHighSpec vs. HighSpec[OneWord], ScalHighSpec vs. HighSpec[OneWord]; Fig. 3). These results are consistent with the hypothesis that early composition should be absent for context-sensitive scalars.

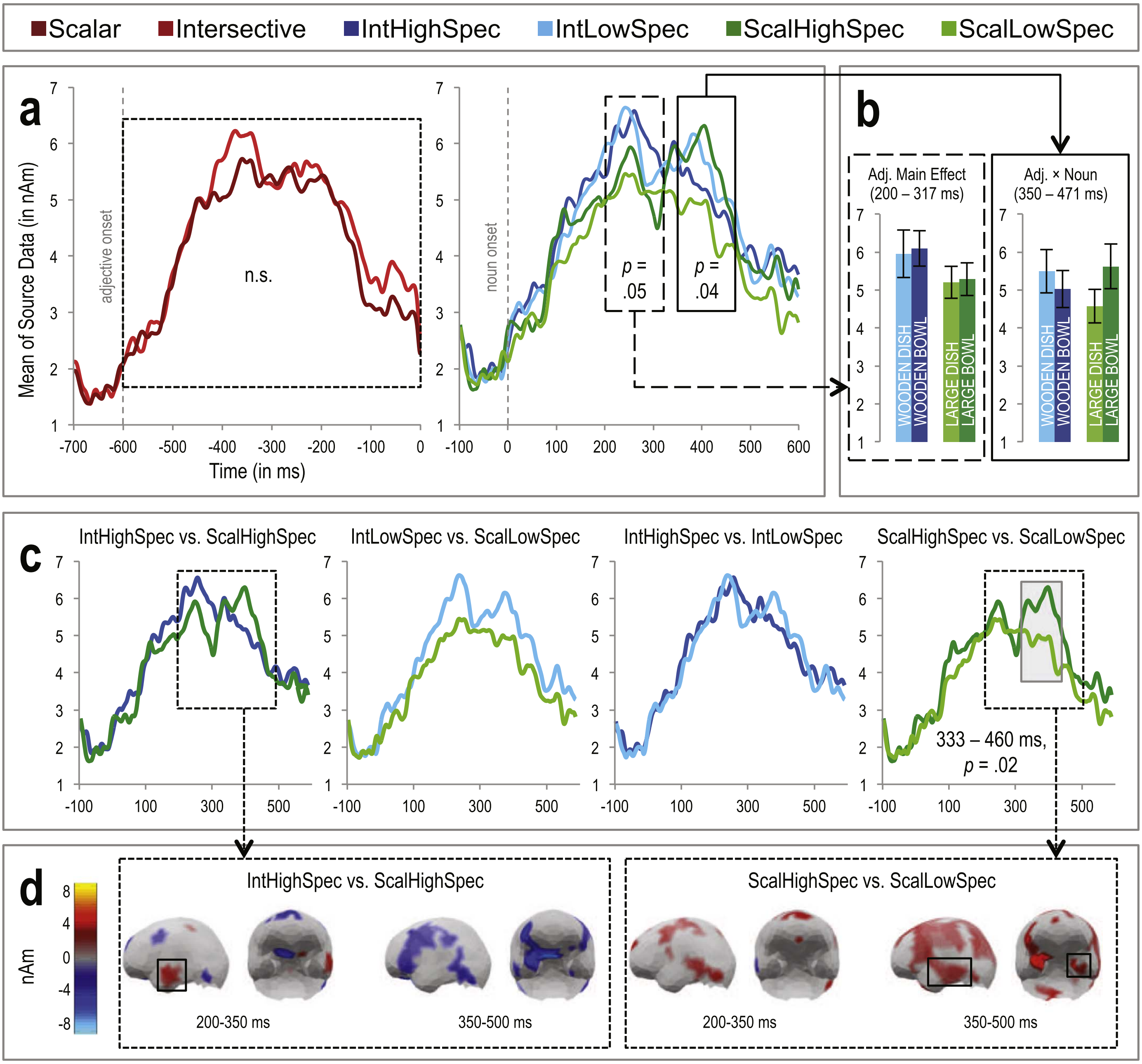

Fig. 2. Results from 2 × 2 ANOVA and corresponding pairwise comparisons, including whole brain plots. (a) Activation (in nAm) by condition in BA 21, where −600 ms represents adjective onset and 0 ms represents noun onset. A 2 × 2 cluster-based permutation ANOVA, time-locked to noun onset, revealed a main effect of Adjective Type between 200 and 317 ms, and an interaction between Adjective Type and Noun Type between 350 and 471 ms (right). No significant differences were identified prior to noun onset (left; time-locked to adjective onset). (b) Bar graphs show mean activity by condition in the two significant clusters identified by the 2 × 2 permutation ANOVA at noun onset. Error bars show SEMs. (c) Pairwise comparisons unpacking early Adjective Type main effect and late Adjective Type by Noun Type interaction. Permutation *t*-tests revealed a qualitatively similar, though non-significant, early pattern of results for Interesective over Scalar adjectives, within both noun types individually (left), and a reliable cluster of increased activity at 333–460 ms for ScalHighSpec over ScalLowSpec (right). (d) Whole brain comparisons at noun onset show activity (in red) if the first condition elicited greater activity than the second condition across at least five adjacent sources and for at least five consecutive milliseconds, at $p < .05$. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Lastly, the effect of noun specificity within the non-combinatory one-word conditions (i.e., HighSpec[OneWord] vs. LowSpec[OneWord]) was examined in a permutation *t*-test over the entire 200–500 ms interval, revealing only a marginally reliable cluster of activity at 200–255 ms (p=.09, one-tailed). However, visual inspection of the waveforms clearly showed that the high-specificity nouns elicited consistently more activity than the low-specificity nouns over the entire epoch following the noun's presentation. Indeed, a post hoc *t*-test on averaged activity over the entire post-noun interval (0–600 ms) was also significant, $t(23)=2.68$, $p=.007$, one-tailed (Fig. 4(a) and (b)).

Despite the fact that our HighSpec and LowSpec nouns were not matched on frequency, we nevertheless found the same general pattern of results as Westerlund and Pylkkänen (2014), whose high-specificity and low-specificity nouns were, in fact, frequency-matched. Thus, frequency likely did not play a role in the pattern of results found in our data, above and beyond that driven by our experimental contrasts of interest.

3.2.2. *Whole brain results*

Uncorrected pairwise full brain comparisons were conducted (1) between the two adjective types separately across the two noun

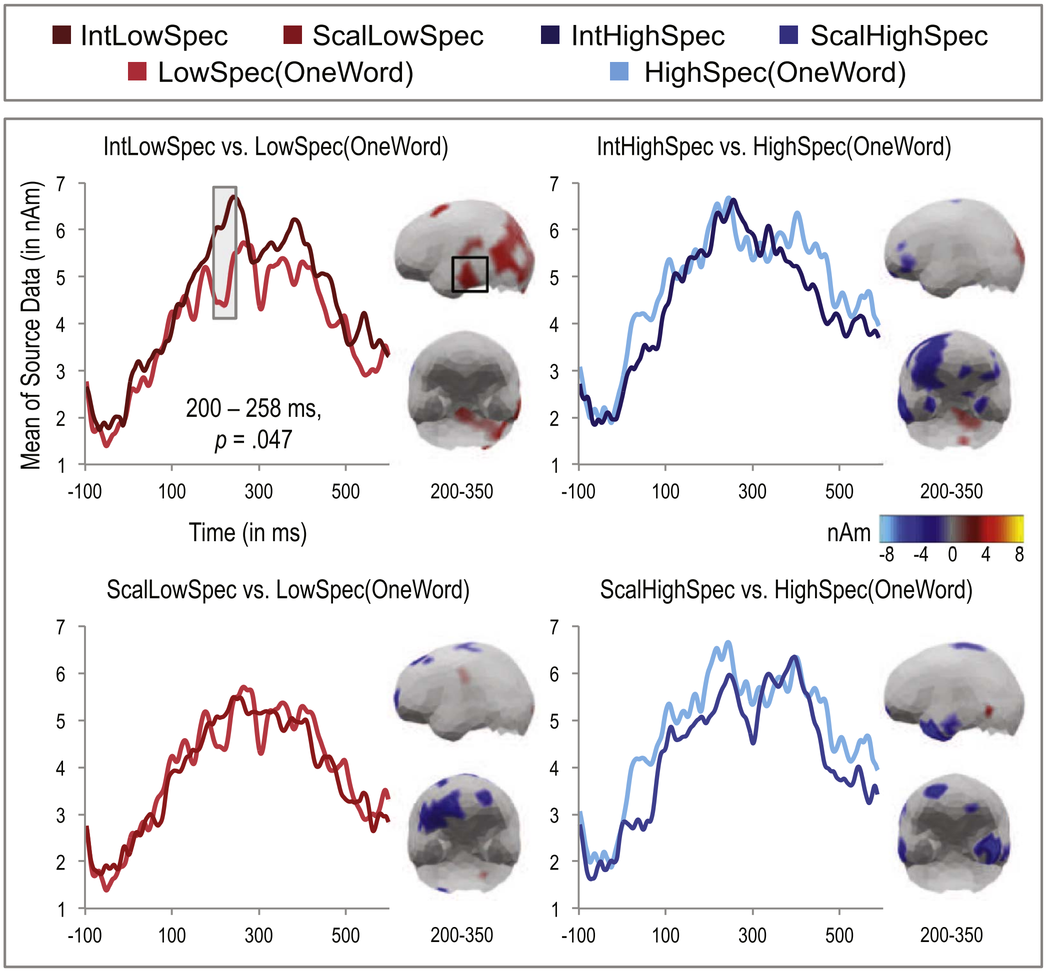

Fig. 3. Pairwise comparisons of combinatory conditions vs. one-word controls (i.e., composition effects), including whole brain plots. Permutation *t*-tests (one-tailed) revealed a reliable cluster of increased activity at 200–258 ms for IntLowSpec over LowSpec (OneWord). Whole brain comparisons show activity (in red) if the first condition elicited greater activity than the second condition across at least five adjacent sources and for at least five consecutive milliseconds, at $p < .05$. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

types (i.e., IntLowSpec vs. ScalLowSpec, IntHighSpec vs. ScalHighSpec), (2) between the two noun types separately across the two adjective types (i.e., IntHighSpec vs. IntLowSpec, ScalHighSpec vs. ScalLowSpec), (3) between each two-word combinatory condition and its one-word non-compositional control (i.e., IntLowSpec vs. LowSpec[OneWord], ScalLowSpec vs. LowSpec[OneWord], IntHighSpec vs. HighSpec[OneWord], ScalHighSpec vs. HighSpec[OneWord]), and (4) between the two non-compositional noun types on their own (i.e., HighSpec[OneWord] vs. LowSpec[OneWord]). Resulting activity was then plotted on the BESA standard brain. For the 2×2 comparisons, we collapsed temporally over an early and late time window (i.e., 200–350 ms, 350–500 ms); for the composition effect comparisons, we collapsed temporally over only an early time window (i.e., 200–350 ms); and for the noun comparison, we collapsed over each 100 ms time interval from 0 to 600 ms post-noun onset.

The results of these comparisons show LATL activation in general concordance with our findings above (see Fig. 2(d) for representative contrasts): A pairwise comparison between intersective and scalar modification of high-specificity nouns (i.e., IntHighSpec vs. ScalHighSpec) showed a clear activity increase for intersectives within the anterior portion of BA 21 in the early time window, and a pairwise comparison of scalar modification with high-specificity vs. low-specificity nouns (i.e., ScalHighSpec vs. ScalLowSpec) revealed an increase in activity for the high-specificity nouns within a somewhat more posterior portion of BA 21, as well as the left temporal pole, in the late time window. However, our early IntHighSpec vs. ScalHighSpec contrast also revealed activity in right ventromedial prefrontal cortex (vmPFC) and the left inferior frontal gyrus (LIFG), although in the opposite direction of the LATL effects, while our later ScalHighSpec vs. ScalLowSpec contrast revealed additional activity in the right temporal pole, as well as the LIFG and left angular gyrus (AG). Several of these regions have also been implicated in semantic combinatory processing across a range of studies and methodologies (LIFG: Hagoort, 2005; Binder et al., 2009; Friederici, 2012; vmPFC: Binder et al., 2009; Bemis and Pylkkänen, 2011; Hagoort, 2013; Pylkkänen et al., 2014; inferior parietal lobule/AG: Lau et al., 2008; Binder et al., 2009; Bemis and Pylkkänen, 2012; right temporal pole: Bemis and Pylkkänen, 2011). Nevertheless, our predictions concerned only the LATL, and we therefore refrain from speculating on the exact nature of these effects.

The composition effect plots revealed increased activation for intersectively-modified low-specificity nouns over their non-compositional one-word controls (i.e., IntLowSpec vs. LowSpec[OneWord]) in the more posterior portion of left BA 21, as well as the left AG, generally consistent with our ROI analysis, while the remaining contrasts included activity in right vmPFC, the left and right temporal poles, and the more superior right frontal lobe, though in the opposite direction (Fig. 3).

Finally, the single-word noun specificity whole brain plots were also largely consistent with our ROI analysis: A pairwise

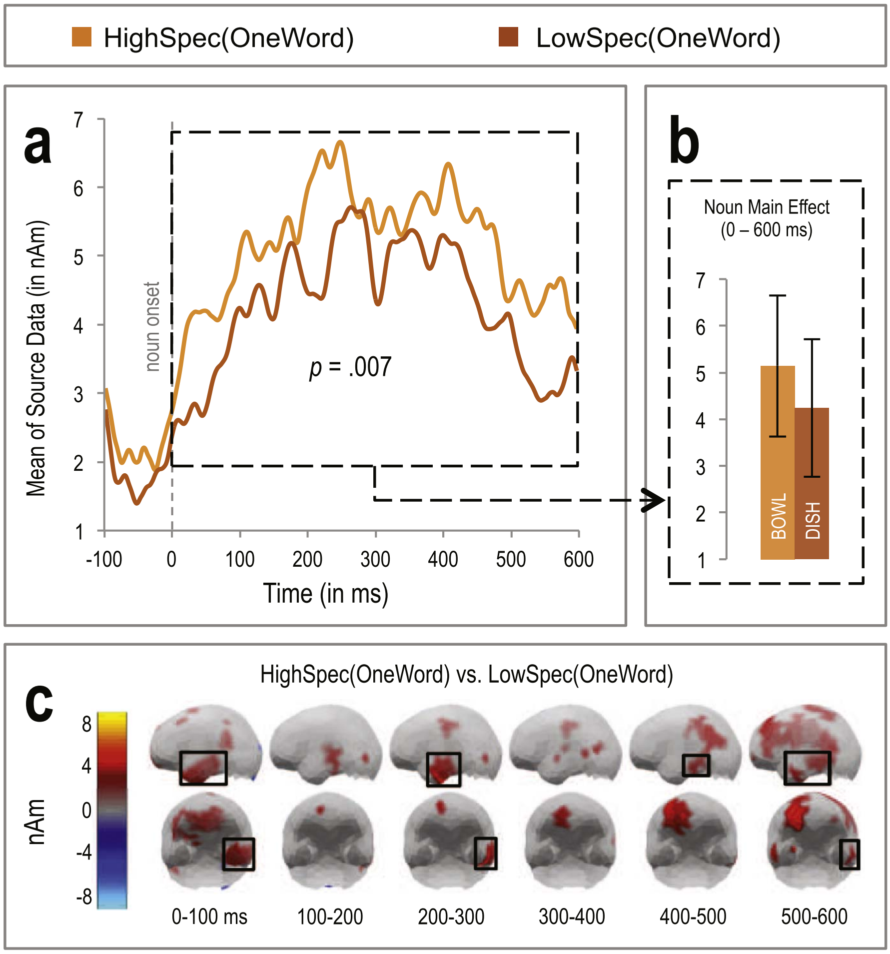

Fig. 4. Effect of noun specificity within non-combinatory one-word conditions. (a) Activation (in nAm) by condition in BA 21, where 0 ms represents noun onset. A *t*-test (one-tailed) on averaged activity over the entire post-noun time interval (0–600 ms) revealed a main effect of Noun Type. (b) Bar graph shows mean activity by condition over the entire post-noun time window. Error bars show SEMs. (c) Whole brain comparisons show activity (in red) if the first condition elicited greater activity than the second condition across at least five adjacent sources and for at least five consecutive milliseconds, at $p < .05$. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

comparison between the two one-word conditions (i.e., HighSpec [OneWord] vs. LowSpec[OneWord]) revealed an increase in activity for the high-specificity over low-specificity nouns in and around left BA 21 and surrounding LATL regions (e.g., left temporal pole) over a large portion of the entire 0–600 ms time window (Fig. 4 (c)). As before, however, these plots also revealed very early right vmPFC activation, as well as relatively late (400–600 ms) activity in the left AG and more superior right frontal lobe.

3.3. Adjective time window

The effect of Adjective Type prior to noun onset was examined over the entire adjective time window (i.e., −600 to 0) in a post hoc permutation *t*-test. No significant clusters were found. Indeed, although visual inspection of the waveforms showed differences between the two conditions prior to noun onset, these differences were not significant, $t(23)=1.07$, $p=.29$ (Fig. 2(a)).

4. Discussion

The primary aim of this work was to better understand the timing of previously reported LATL composition effects (Bemis and Pylkkänen, 2011, 2012, 2013; Del Prato and Pylkkänen, 2014; Pylkkänen et al., 2014; Westerlund and Pylkkänen, 2014; Westerlund et al., 2015), especially in the context of the extreme variability in the reported timeframes of lexical-semantic access to date (e.g., Kutas and Van Petten, 1988; Kutas and Federmeier, 2000; Pulvermüller et al., 2001; Pylkkänen and Marantz, 2003; Shtyrov et al., 2004; Pulvermüller, 2005; Pulvermüller et al., 2005; Van Petten and Luka, 2006; Federmeier, 2007; Lau et al., 2008, 2009; Kutas and Federmeier, 2009, 2011). To do so, we manipulated both the type of adjective and the type of noun as the input to simple adjective-noun phrases, modeled on the paradigm employed in a series of recent MEG studies (Bemis and Pylkkänen, 2011, 2012; Westerlund and Pylkkänen, 2014; Westerlund et al., 2015). Our adjective manipulation included both scalar and intersective adjectives, which differ from each other in context-sensitivity, and our nouns were conceptually either more or less specific, following Westerlund and Pylkkänen (2014). The adjective manipulation was intended to test whether LATL combinatory effects may reflect an early initial stage of composition, at which only intersectives could compose; and the noun manipulation, to reveal a potentially later stage of composition, at which the size of the comparison class provided for a context-sensitive scalar adjective might matter.

Conforming to these hypotheses, our combinatory phrases exhibited a main effect of Adjective Type in an early time window, with nouns in intersective contexts eliciting greater left temporal lobe activation than those following scalar adjectives. Importantly, no reliable differences between the two adjective types were observed prior to noun onset. If semantic activation is gradient and intersectives are capable of combining with more rudimentary noun meanings than scalars, this pattern is straightforwardly explained. Notably, however, the early composition effect was further limited to nouns with more general meanings, as also observed in Westerlund and Pylkkänen (2014). The picture that is emerging, then, suggests that early composition in the LATL only occurs when two relatively "easy-to-process" items compose, such as, in this case, context-independent adjectives and conceptually rather general nouns, with, by hypothesis, fewer features to activate.

One way to conceptualize this within a broader context is that the LATL constitutes an early node within the processing path of words, with the signal arriving there shortly after the activation of visual word form regions along the ventral surface of temporal cortex, a hypothesis straightforwardly supported by prior MEG studies on the spatio-temporal dynamics of word processing (e.g., Marinkovic et al., 2003).[1] Importantly, studies on single word processing show that the LATL does not "specialize" in composition, but rather is activated by all words, whether or not they are in a combinatory context.[2] However, when the incoming word does occur in a combinatory context, an increase in LATL activity can be elicited, as shown by the studies forming the background for the current work (Bemis and Pylkkänen, 2011, and subsequent findings). Thus, one hypothesis consistent with the data available to date is that the early LATL activity serves to activate *and* build initial "drafts" of conceptual representations. According to the present results, prerequisites for a current item Y to combine with a context item X in the LATL (to form a phrase [X Y]) are two-fold: (a) the meaning of Y needs to be sufficiently simple such that it can at least to some extent be activated by 200 ms, and (b) X needs to not require a particularly well fleshed-out semantic representation of Y in order to combine with it. Under this hypothesis, the explanation of the current findings is that the scalar modifiers fail to meet (b), whereas the high-specificity nouns fail to meet (a), leaving only the intersective + low-specificity pairings to elicit a combinatory effect. Accordingly, one way to think of the early combinatory role of the LATL is that it is an opportunistic combiner, composing meanings when it can; when it cannot, however, the signal passes through it uncomposed. Uncomposed single-word meanings may still show semantic effects, though at least in our studies they have been statistically weaker than the combinatory effects (Westerlund and Pylkkänen, 2014; Zhang and Pylkkänen, 2015), including the currently reported specificity effect. These results suggest that the dynamic building of new meanings may be a somewhat stronger activator of the LATL than access to already stored single-word meanings.

Once the signal has left the LATL by roughly 300 ms, does it ever return there? Our data provide some tentative evidence that it does. Specifically, we observed a more complicated later pattern of LATL activity, conforming to the type of sensitivity to noun specificity within the scalar items that originally motivated our specificity manipulation—i.e., the size of the comparison class provided by the noun mattered for scalar modifiers in a later processing stage. Specifically, when scalar adjectives combined with nouns describing narrower (i.e., more specific) comparison classes, more left temporal lobe activity was observed as compared to nouns with more general meanings. This could be interpreted as a reflection of the fact that when scalar modifiers are combined with nouns with less specific meanings, the meanings of the scalars themselves remain rather vague. For example, the set of *large animals* includes whales, elephants, giraffes, rhinoceroses, and many other types of animals that vary wildly in size from one to another. In contrast, the set of *large dogs* is smaller and less variable, making the adjective *large* more contrasting and "diagnostic" (e.g., Smith and Osherson, 1984; Smith et al., 1988) in the context of dogs than in the context of animals more generally.

Conversely, noun specificity did not affect the LATL amplitudes of intersective phrases in this later time window, a result that could be predicted both by the fact that the notion of a comparison class is irrelevant for the processing of intersectives, and by the hypothesis that the processing of these phrases may in this later time window simply be over. Overall, though, the evidence for any late combinatory stage must remain quite preliminary, given that we did not directly test for whether scalar phrases elicited larger amplitudes than their one-word controls in this later time window. We leave this as an open question for future work.

Crucially, as already emphasized in the Introduction, our findings relate solely to the LATL, to the exclusion of the many other brain regions often also implicated in linguistic composition. Specifically, we set forth to address the question: Given the LATL's apparent involvement in semantic composition, how can we reconcile the early timing of this effect with the rich body of literature placing lexical-semantic access at a much, though not always, later point in time? Thus, although this work suggests that early composition in the LATL occurs only for those semantic representations that are shallowest or most easily accessible, it remains an open question as to how, if at all, the LATL's linguistic computational role is different from those of other brain regions, which may or may not turn out to exhibit similarly gradient response profiles.

5. Conclusion

In sum, this study investigated the nature of the LATL's combinatory role in language processing by taking advantage of the context-sensitivity of scalar as compared to intersective adjectives in simple adjective-noun phrases. In general, our results suggest that linguistic composition in the left temporal lobe unfolds over time, with previously documented effects at roughly 200–300 ms post-stimulus onset representing a more initial stage of semantic processing, at which time only the most readily available semantic representations are accessed and composed. The interpretations computed at this early stage may then change over time as additional representational detail subsequently becomes available. The current results suggest a second stage of semantic processing at approximately 350–450 ms, during which LATL amplitudes increased when scalar adjectives were provided with narrower comparison classes.

[1] Interestingly, a similar timing of LATL activation (~200–250 ms) is also observed in production tasks using picture naming, as evidenced both by electrocorticogram data on single word production (Chen et al., 2016) and by MEG data on phrase production (Pylkkänen et al., 2014; Del Prato and Pylkkänen, 2014; Blanco-Elorrieta and Pylkkänen, 2016). The same holds for the auditory comprehension of single words (Marinkovic et al., 2003). Combinatory effects for auditory phrases have been reported to onset at around 250 ms—i.e., slightly later than parallel effects in the visual modality (Bemis and Pylkkänen, 2013). Thus, although the gradual unfolding of auditory words may delay LATL effects at least in some circumstances, overall, it appears that the timing of LATL activation is relatively fixed in the 200–250 ms time window, no matter the input.

[2] This is consistent with the hypothesis put forth by Blank et al. (2016) that combinatory brain regions are in general the same as lexical regions, though one possible explanation of this could be that single words always activate their possible contexts, as shown by both behavioral studies (McDonald and Shillcock, 2001; Baayen et al., 2011) and MEG research on the LATL (Linzen et al., 2013). An alternative account could of course be that the single-word and combinatory activations in the LATL occur in somewhat different locations, a hypothesis that cannot be ruled out on the basis of the currently available data.

Acknowledgments

This research was funded by the National Science Foundation grant BCS-1221723 (L.P.) and grant G1001 from the NYUAD Institute, New York University Abu Dhabi (L.P.). We thank Chris Barker, Alec Marantz, the members of NYU's Neuroscience of Language Lab, and the anonymous reviewers for valuable comments and feedback; Jeffrey Walker for MEG technical support; and especially Paul Del Prato for his expertise with data preprocessing and analysis.

References

Aslin, R.N., Saffran, J.R., Newport, E.L., 1998. Computation of conditional probability statistics by 8-month-old infants. Psychol. Sci. 9 (4), 321–324.
Baayen, R.H., Milin, P., Đurđević, D.F., Hendrix, P., Marelli, M., 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. Psychol. Rev. 118 (3), 438–481.
Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., Treiman, R., 2007. The English Lexicon Project. Behav. Res. 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. http://dx.doi.org/10.1016/j.neuroimage.2011.01.066.
Baron, S.G., Thompson-Schill, S., Weber, M., Osherson, D., 2010. An early stage of conceptual combination: superimposition of constituent concepts in left anterolateral temporal lobe. Cogn. Neurosci. 1 (1), 44–51. http://dx.doi.org/10.1080/17588920903548751.
Bates, D.M., 2010. lme4: Mixed-effects Modeling with R. Springer, New York http://lme4.r-forge.r-project.org/book.
Bemis, D.K., Pylkkänen, L., 2011. Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. J. Neurosci. 31 (8), 2801–2814. http://dx.doi.org/10.1523/JNEUROSCI.5003-10.2011.
Bemis, D.K., Pylkkänen, L., 2012. Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. Cereb. Cortex 23 (8), 1859–1873. http://dx.doi.org/10.1093/cercor/bhs170.
Bemis, D.K., Pylkkänen, L., 2013. Flexible composition: MEG evidence for the deployment of basic combinatorial linguistic mechanisms in response to task demands. PLoS One 8 (9), e73949. http://dx.doi.org/10.1371/journal.pone.0073949.
Bierwisch, M., 1989. The semantics of gradation. In: Bierwisch, M., Lang, E. (Eds.), Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation. Springer, Berlin, pp. 71–261.
Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends Cogn. Sci. 15 (11), 527–536. http://dx.doi.org/10.1016/j.tics.2011.10.001.
Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L., 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb. Cortex 19 (12), 2767–2796. http://dx.doi.org/10.1093/cercor/bhp055.
Blanco-Elorrieta, E., Pylkkänen, L., 2016. Composition of complex numbers: delineating the computational role of the left anterior temporal lobe. NeuroImage 124, 194–203. http://dx.doi.org/10.1016/j.neuroimage.2015.08.049.
Blank, I., Balewski, Z., Mahowald, K., Fedorenko, E., 2016. Syntactic processing is distributed across the language system. NeuroImage 127, 307–323. http://dx.doi.org/10.1016/j.neuroimage.2015.11.069.
Bonner, M.F., Peelle, J.E., Cook, P.A., Grossman, M., 2013. Heteromodal conceptual processing in the angular gyrus. NeuroImage 71, 175–186. http://dx.doi.org/10.1016/j.neuroimage.2013.01.006.
Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10 (4), 433–436.
Brennan, J., Pylkkänen, L., 2012. The time-course and spatial distribution of brain activity associated with sentence processing. NeuroImage 60 (2), 1139–1148. http://dx.doi.org/10.1016/j.neuroimage.2012.01.030.
Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D.J., Pylkkänen, L., 2010. Syntactic structure building in the anterior temporal lobe during natural story listening. Brain Lang. 120 (2), 163–173. http://dx.doi.org/10.1016/j.bandl.2010.04.002.
Bright, P., Moss, H., Tyler, L.K., 2004. Unitary vs. multiple semantics: PET studies of word and picture processing. Brain Lang. 89 (3), 417–432.
Chen, Y., Shimotake, A., Matsumoto, R., Kunieda, T., Kikuchi, T., Miyamoto, S., Lambon Ralph, M.A., 2016. The "when" and "where" of semantic coding in the anterior temporal lobe: temporal representational similarity analysis of electrocorticogram data. Cortex 79, 1–13.
Clarke, A., Taylor, K.I., Tyler, L.K., 2011. The evolution of meaning: spatio-temporal dynamics of visual object recognition. J. Cogn. Neurosci. 23 (8), 1887–1899. http://dx.doi.org/10.1162/jocn.2010.21544.
Clarke, A., Taylor, K.I., Devereux, B., Randall, B., Tyler, L.K., 2013. From perception to conception: how meaningful objects are processed over time. Cereb. Cortex 23 (1), 187–197. http://dx.doi.org/10.1093/cercor/bhs002.
Davies, M., 2011. Google Books (American English) Corpus (155 Billion Words, 1810–2009). Available online at: ⟨http://googlebooks.byu.edu⟩.
Del Prato, P., Pylkkänen, L., 2014. MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. Front. Psychol. 5, 524. http://dx.doi.org/10.3389/fpsyg.2014.00524.
Endrass, T., Mohr, B., Pulvermüller, F., 2004. Enhanced mismatch negativity brain response after binaural word presentation. Eur. J. Neurosci. 19 (6), 1653–1660.
Federmeier, K.D., 2007. Thinking ahead: the role and roots of prediction in language comprehension. Psychophysiology 44 (4), 491–505.
Federmeier, K.D., Kutas, M., 1999. A rose by any other name: long-term memory structure and sentence processing. J. Mem. Lang. 41 (4), 469–495.
Fellbaum, C., 1998. WordNet: an Electronic Lexical Database. MIT Press, Cambridge http://wordnet.princeton.edu.
Fischler, I., Bloom, P.A., Childers, D.G., Roucos, S.E., Perry Jr., N.W., 1983. Brain potentials related to stages of sentence verification. Psychophysiology 20 (4), 400–409.
Friederici, A.D., 2012. The cortical language circuit: from auditory perception to sentence comprehension. Trends Cogn. Sci. 16 (5), 262–268. http://dx.doi.org/10.1016/j.tics.2012.04.001.
Friederici, A.D., Meyer, M., Von Cramon, D.Y., 2000. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang. 74 (2), 289–300. http://dx.doi.org/10.1006/brln.2000.2313.
Gainotti, G., 2006. Anatomical functional and cognitive determinants of semantic memory disorders. Neurosci. Biobehav. Rev. 30 (5), 577–594.
Gainotti, G., 2007. Different patterns of famous people recognition disorders in patients with right and left anterior temporal lesions: a systematic review. Neuropsychologia 45 (8), 1591–1607.
Gainotti, G., 2012. The format of conceptual representations disrupted in semantic dementia: a position paper. Cortex 48 (5), 521–529. http://dx.doi.org/10.1016/j.cortex.2011.06.019.
Garrard, P., Carroll, E., 2006. Lost in semantic space: a multi-modal, non-verbal assessment of feature knowledge in semantic dementia. Brain 129 (5), 1152–1163.
Garrard, P., Hodges, J.R., 2000. Semantic dementia: clinical, radiological, and pathological perspectives. J. Neurol. 247 (6), 409–422.
Gauthier, I., Anderson, A.W., Tarr, M.J., Skudlarski, P., Gore, J.C., 1997. Levels of categorization in visual recognition studied using functional magnetic resonance imaging. Curr. Biol. 7 (9), 645–651.
Gorno-Tempini, M.L., Rankin, K.P., Woolley, J.D., Rosen, H.J., Phengrasamy, L., Miller, B.L., 2004. Cognitive and behavioral profile in a case of right anterior temporal lobe neurodegeneration. Cortex 40 (4–5), 631–644.
Grabowski, T.J., Damasio, H., Tranel, D., Ponto, L.L.B., Hichwa, R.D., Damasio, A.R., 2001. A role for left temporal pole in the retrieval of words for unique entities. Hum. Brain Mapp. 13 (4), 199–212.
Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9 (9), 416–423.
Hagoort, P., 2013. MUC (Memory, Unification, Control) and beyond. Front. Psychol. 4, 416.
Hämäläinen, M.S., Ilmoniemi, R.J., 1994. Interpreting magnetic fields of the brain: minimum norm estimates. Med. Biol. Eng. Comput. 32 (1), 35–42.
Hinojosa, J.A., Martin-Loeches, M., Muñoz, F., Casado, P., Pozo, M.A., 2004. Electrophysiological evidence of automatic early semantic processing. Brain Lang. 88 (1), 39–46.
Hodges, J.R., Graham, N., Patterson, K., 1995. Charting the progression in semantic dementia: implications for the organization of semantic memory. Memory 3 (3–4), 463–495.
Hodges, J.R., Patterson, K., Oxbury, S., Funnell, E., 1992. Semantic dementia: progressive fluent aphasia with temporal lobe atrophy. Brain 115 (6), 1783–1806.
Holcomb, P.J., McPherson, W.B., 1994. Event-related brain potentials reflect semantic priming in an object decision task. Brain Cogn. 24 (2), 259–276.
Humphries, C., Binder, J.R., Medler, D.A., Liebenthal, E., 2006. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. J. Cogn. Neurosci. 18 (4), 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 (3), 924–932.
Humphries, C., Love, T., Swinney, D., Hickock, G., 2005. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. Hum. Brain Mapp. 26 (2), 128–138. http://dx.doi.org/10.1002/hbm.20148.
Kamp, H., 1975. Two theories about adjectives. In: Keenan, E. (Ed.), Formal Semantics of Natural Language. Cambridge University Press, Cambridge, pp. 123–155.
Kamp, H., Partee, B., 1995. Prototype theory and compositionality. Cognition 57 (2), 129–191.
Kennedy, C., 1999. Projecting the Adjective: the Syntax and Semantics of Gradability and Comparison. Garland Press, New York.
Kennedy, C., 2007. Vagueness and grammar: the semantics of relative and absolute gradable predicates. Linguist. Philos. 30 (1), 1–45. http://dx.doi.org/10.1007/s10988-006-9008-0.
Kennedy, C., 2012. Adjectives. In: Russell, G., Graff Fara, D. (Eds.), Routledge Companion to Philosophy of Language. Routledge, New York.
Kennedy, C., McNally, L., 2005. Scale structure and the semantic typology of gradable predicates. Language 81 (2), 345–381.
Klein, E., 1980. A semantics for positive and comparative adjectives. Linguist. Philos. 4 (1), 1–45. http://dx.doi.org/10.1007/BF00351812.
Kleiner, M., Brainard, D.H., Pelli, D., Ingling, A., Murray, R., Broussard, C., 2007. What's new in Psychtoolbox-3. Perception 36 (14), 1–16.
Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn. Sci. 4 (12), 463–470.
Kutas, M., Federmeier, K.D., 2009. N400. Scholarpedia 4 (10), 7790.

Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.
Kutas, M., Van Petten, C., 1988. Event-related brain potential studies of language. In: Ackles, P.K., Jennings, J.R., Coles, M.G.H. (Eds.), Advances in Psychophysiology 3. JAI Press, Inc., Greenwich, pp. 139–187.
Kutas, M., Van Petten, C., 1994. Psycholinguistics electrified: event-related brain potential investigations. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, San Diego, pp. 83–143.
Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Fox, P. T., 2000. Automated Talairach atlas labels for functional brain mapping. Hum. Brain Mapp. 10 (3), 120–131.
Landauer, T.K., Foltz, P.W., Laham, D., 1998. Introduction to latent semantic analysis. Discourse Process. 25 (2–3), 259–284. http://dx.doi.org/10.1080/01638539809545028.
Laszlo, S., Federmeier, K.D., 2010. The N400 as a snapshot of interactive processing: evidence from regression analyses of orthographic neighbor and lexical associate effects. Psychophysiology 48 (2), 176–186. http://dx.doi.org/10.1111/j.1469–8986.2010.01058.x.
Lau, E.F., Almeida, D., Hines, P.C., Poeppel, D., 2009. A lexical basis for N400 context effects: evidence from MEG. Brain Lang. 111 (3), 161–172. http://dx.doi.org/10.1016/j.bandl.2009.08.007.
Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (De)constructing the N400. Nat. Rev. Neurosci. 9 (12), 920–933. http://dx.doi.org/10.1038/nrn2532.
Leffel, T., Lauter, M., Westerlund, M., Pylkkänen, L., 2014. Restrictive vs. non-restrictive composition: a magnetoencephalography study. Lang. Cogn. Neurosci. 29 (10), 1191–1204.
Linzen, T., Marantz, A., Pylkkänen, L., 2013. Syntactic context effects in visual word recognition: an MEG study. Ment. Lex. 8 (2), 117–139. http://dx.doi.org/10.1075/ml.8.2.01lin.
Marinkovic, K., Dhond, R.P., Dale, A.M., Glessner, M., Carr, V., Halgren, E., 2003. Spatiotemporal dynamics of modality-specific and supramodal word processing. Neuron 38 (3), 487–497.
Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164 (1), 177–190.
Marslen-Wilson, W.D., 1973. Linguistic structure and speech shadowing at very short latencies. Nature 244 (5417), 522–523.
Marslen-Wilson, W.D., 1985. Speech shadowing and speech comprehension. Speech Commun. 4, 55–73.
Marslen-Wilson, W.D., 1987. Functional parallelism in spoken word-recognition. Cognition 25 (1–2), 71–102.
Marslen-Wilson, W.D., Tyler, L.K., 1975. Processing structure of sentence perception. Nature 257 (5529), 784–786.
Mazoyer, B.M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Mehler, J., 1993. The cortical representation of speech. J. Cogn. Neurosci. 5 (4), 467–479. http://dx.doi.org/10.1162/jocn.1993.5.4.467.
McDonald, S.A., Shillcock, R.C., 2001. Rethinking the word frequency effect: the neglected role of distributional information in lexical processing. Lang. Speech 44 (3), 295–323.
Miller, G.A., Selfridge, J.A., 1950. Verbal context and the recall of meaningful material. Am. J. Psychol. 63 (2), 176–185.
Mohr, B., Pulvermüller, F., 2002. Redundancy gains and costs in cognitive processing: the effects of short stimulus onset asynchronies. J. Exp. Psychol.: Learn. Mem. Cogn. 28 (6), 1200–1223.
Moss, H.E., 1997. The time course of activation of semantic information during spoken word recognition. Lang. Cogn. Process. 12 (5/6), 695–731.
Mummery, C.J., Patterson, K., Price, C.J., Ashburner, J., Frackowiak, R.S., Hodges, J.R., 2000. A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. Ann. Neurol. 47 (1), 36–45.
Mummery, C.J., Patterson, K., Wise, R.J., Vandenbergh, R., Price, C.J., Hodges, J.R., 1999. Disrupted temporal lobe connections in semantic dementia. Brain 122 (1), 61–73.
Pallier, C., Devauchelle, A.D., Dehaene, S., 2011. Cortical representation of the constituent structure of sentences. Proc. Natl. Acad. Sci. 108 (6), 2522–2527. http://dx.doi.org/10.1073/pnas.1018711108.
Partee, B., 1995. Lexical semantics and compositionally. In: Gleitman, L., Liberman, M. (Eds.), An Invitation to Cognitive Science: Language. MIT Press, Cambridge, pp. 311–360.
Patterson, K., Lambon Ralph, M.A., Jefferies, E., Woollams, A., Jones, R., Hodges, J.R., Rogers, T.T., 2006. "Presemantic" cognition in semantic dementia: six deficits in search of an explanation. J. Cogn. Neurosci. 18 (2), 169–183.
Pobric, G., Jefferies, E., Lambon Ralph, M.A., 2007. Anterior temporal lobes mediate semantic representation: mimicking semantic dementia by using rTMS in normal participants. Proc. Natl. Acad. Sci. 104 (50), 20137–20141.
Pulvermüller, F., 2005. Brain mechanisms linking language and action. Nat. Rev. Neurosci. 6 (7), 576–582.
Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Näätänen, R., 2001. Memory traces for words as revealed by the mismatch negativity. NeuroImage 14 (3), 607–616.
Pulvermüller, F., Shtyrov, Y., Ilmoniemi, R., 2005. Brain signatures of meaning access in action word recognition. J. Cogn. Neurosci. 17 (6), 884–892.
Pylkkänen, L., 2015. Composition of complex meaning: interdisciplinary perspectives on the left anterior temporal lobe. In: Hickok, G., Small, S.L. (Eds.), Neurobiology of Language. Elsevier, New York.
Pylkkänen, L., Bemis, D.K., Blanco Elorrieta, E., 2014. Building phrases in language production: an EMG study of simple composition. Cognition 133 (2), 371–384. http://dx.doi.org/10.1016/j.cognition.2014.07.001.
Pylkkänen, L., Marantz, A., 2003. Tracking the time course of word recognition with MEG. Trends Cogn. Sci. 7 (5), 187–189.
Pylkkänen, L., Stringfellow, A., Marantz, A., 2002. Neuromagnetic evidence for the timing of lexical activation: an MEG component sensitive to phonotactic probability but not to neighborhood density. Brain Lang. 81 (1–3), 666–678.
Rastle, K., Davis, M.H., Marslen-Wilson, W.D., Tyler, L.K., 2000. Morphological and semantic effects in visual word recognition: a time-course study. Lang. Cogn. Process. 15 (4/5), 507–537.
Rodd, J.M., Davis, M.H., Johnsrude, I.S., 2005. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. Cereb. Cortex 15 (8), 1261–1269.
Rogalsky, C., Hickock, G., 2009. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. Cereb. Cortex 19 (4), 786–796. http://dx.doi.org/10.1093/cercor/bhn126.
Rogers, T.T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M.L., Patterson, K., Price, C.J., 2006. Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. Cogn. Affect. Behav. Neurosci. 6 (3), 201–213.
Rogers, T.T., Lambon Ralph, M.A., Garrard, P., Bozeat, S., McClelland, J.L., Hodges, J.R., Patterson, K., 2004. Structure and deterioration of semantic memory: a neuropsychological and computational investigation. Psychol. Rev. 111 (1), 205–235.
Rugg, M.D., 1990. Event-related brain potentials dissociate repetition effects of high- and low-frequency words. Mem. Cogn. 18 (4), 367–379.
Schwartz, M.F., Marin, O.S., Saffran, E.M., 1979. Dissociations of language function in dementia: a case study. Brain Lang. 7 (3), 277–306.
Shimotake, A., Matsumoto, R., Ueno, T., Kunieda, T., Saito, S., Hoffman, P., Lambon Ralph, M.A., 2015. Direct exploration of the role of the ventral anterior temporal lobe in semantic memory: cortical stimulation and local field potential evidence from subdural grid electrodes. Cereb. Cortex 25 (10), 3802–3817.
Shtyrov, Y., Pulvermüller, F., 2002. Neurophysiological evidence of memory traces for words in the human brain. NeuroReport 13 (4), 521–525.
Shtyrov, Y., Hauk, O., Pulvermüller, F., 2004. Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. Eur. J. Neurosci. 19 (4), 1083–1092.
Shtyrov, Y., Pihko, E., Pulvermüller, F., 2005. Determinants of dominance: Is language laterality explained by physical or linguistic features of speech? NeuroImage 27 (1), 37–47.
Siegel, M.E.A., 1976. Capturing the Adjective (Doctoral dissertation). University of Massachusetts, Amherst http://scholarworks.umass.edu/dissertations/AAI7622300.
Simon, D.A., Lewis, G., Marantz, A., 2012. Disambiguating form and lexical frequency effects in MEG responses using homonyms. Lang. Cogn. Process. 27 (2), 275–287.
Smith, M.E., Halgren, E., 1987. Event-related potentials during lexical decision: effects of repetition, word frequency, pronounceability, and concreteness. Electroencephalogr. Clin. Neurophysiol. (Suppl. 40), S417–S421.
Smith, E.E., Osherson, D., 1984. Conceptual combination with prototype concepts. Cogn. Sci. 8 (4), 337–361.
Smith, E.E., Osherson, D., Rips, L.J., Keane, M., 1988. Combining prototypes: a selective modification model. Cogn. Sci. 12 (4), 485–527.
Snowden, J.S., Goulding, P.J., Neary, D., 1989. Semantic dementia: a form of circumscribed cerebral atrophy. Behav. Neurol. 2 (3), 167–182.
Stowe, L.A., Broere, C.A., Paans, A.M., Wijers, A.A., Mulder, G., Vaalburg, W., Zwarts, F., 1998. Localizing components of a complex task: sentence processing and working memory. NeuroReport 9 (13), 2995–2999.
Tarkiainen, A., Helenius, P., Hansen, P.C., Cornelissen, P.L., Salmelin, R., 1999. Dynamics of letter string perception in the human occipitotemporal cortex. Brain 22 (11), 2119–2132.
Thompson, S.P., Newport, E.L., 2007. Statistical learning of syntax: the role of transitional probability. Lang. Learn. Dev. 3 (1), 1–42.
Tyler, L.K., Stamatakis, E.A., Bright, P., Acres, K., Abdallah, S., Rodd, J.M., Moss, H.E., 2004. Processing objects at different levels of specificity. J. Cogn. Neurosci. 16 (3), 351–362.
Van Petten, C., Luka, B.J., 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain Lang. 97 (3), 279–293.
Westerlund, M., Kastner, I., Al Kaabi, M., Pylkkänen, L., 2015. The LATL as locus of composition: MEG evidence from English and Arabic. Brain Lang. 141, 124–134. http://dx.doi.org/10.1016/j.bandl.2014.12.003.
Westerlund, M., Pylkkänen, L., 2014. The role of the left anterior temporal lobe in semantic composition vs. semantic memory. Neuropsychologia 57, 59–70. http://dx.doi.org/10.1016/j.neuropsychologia.2014.03.001.
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.