CrossMark

The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study

Linmin Zhang [a,*], Liina Pylkkänen [a,b,c]

[a] Department of Linguistics, New York University, 10 Washington Place, New York, NY 10003, USA
[b] Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA
[c] NYUAD Institute, New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE

ARTICLE INFO

ABSTRACT

A fundamental function of natural language is to focus the interlocutor's attention to specific entities and circumstances from the vast set of possibilities in the environment. In other words, as an utterance progresses, the narrower its reference typically becomes. Intriguingly, there is substantial convergence in the neural localization of conceptual specificity effects at the single word level and combinatory effects at the phrasal level, both systematically affecting the left anterior temporal lobe (LATL). However, the relationship between these two types of effects is not well understood. The current study used MEG to characterize the temporal progression of both types of effects in minimal two-word phrases (e.g., *tomato soup*), where single word specificity was varied in both first and second position (e.g., *tomato* vs. *vegetable*; *soup* vs. *dish*). These combinatory phrases were further compared to non-combinatory single nouns of high and low specificity. Our most robust result was an effect of the specificity of the first word while processing the second word: responses to the second word were the largest when it was being composed with a more specific as opposed to a more general modifier. In the modifier position, specificity had no reliable effects, while non-combinatory single nouns did show a subtle LATL increase when specific. In all, our findings show that when non-semantic factors such as frequency are controlled for, conceptual specificity weakly modulates LATL activity in non-combinatory situations (i.e., at a single noun), but robustly affects the size of the LATL composition effect. Thus LATL activity appears to be most strongly driven by the composition of concepts as opposed to access to single concepts.

Introduction

Conceptual access vs. conceptual combination in the left anterior temporal lobe (LATL)

Language provides the human brain a vehicle for the expression of infinitely many and infinitely complex concepts. Some rich concepts can be expressed with single words: for example in English, the concept of an open rectangular handbag that is larger than a purse, typically has two handles and is often made out of canvas can be expressed with just the single syllable *tote*. But the true power of human language lies in its potential to create new concepts via combining existing lexicalized concepts into novel phrases, sentences and narratives. How does the dynamic act of composing new meaning relate to the representation of complex meaning within single words? This question has risen as a core problem for characterizing the brain basis of semantic processing due to the finding that both research on single concept representation (Warrington, 1975; Hodges et al., 1992, 1995; Mummery et al., 1999a,

2000; Rogers et al., 2006; Rogers & Patterson, 2007; Desgranges et al., 2007; Lambon Ralph et al., 2009; Binney et al., 2010) and research on combinatory semantics (Mazoyer et al., 1993; Stowe et al., 1998; Humphries et al., 2001, 2005; Vandenberghe et al., 2002; Friederici & Kotz, 2003; Brennan et al., 2012; Brennan & Pylkkänen, 2012; Bemis & Pylkkänen, 2011, 2013a,b; Westerlund & Pylkkänen, 2014) have converged on the same region as pivotal in both processes: the left anterior temporal lobe (LATL). Does the LATL encode a unified computation of meaning specification, operating both within and across words? Or are the two effects computationally separate, despite their spatial proximity?

The spatiotemporal interaction of single word specificity and composition was recently addressed using MEG (Westerlund & Pylkkänen, 2014), with results showing that when a variety of lexical level factors are controlled for, single word specificity effects are actually rather subtle in the LATL, but in combinatory contexts, specificity robustly modulates the size of the composition effect: less specific nouns, such as *boat*, elicited a clear LATL increase when modified by adjectives (*blue boat* vs. *qwtp boat*) whereas no such effect was reliably observed for more specific nouns (*blue canoe* vs. *qwtp canoe*). Thus it appeared that adjectival modification robustly boosted LATL activity only when the modified nouns were themselves somewhat vague in meaning. This interaction

demonstrates that specificity and composition both affect the same neural activity and offers initial evidence that the construction of complex meanings may modulate LATL activity more strongly than the conceptual specificity of an already stored item (such as *boat* vs. *canoe*).

As a stronger test of whether LATL activity primarily reflects access to stored conceptual representations or the composition of complex meanings, the current MEG study varied not only the conceptual specificity of modified nouns but also the conceptual specificity of the modifiers. With this design we aimed for a maximal chance to observe single word specificity effects both in the first and second position of the phrases as well as to evaluate how specificity in first and second position affects the composition effect. Instead of adjectival modification, we used noun-noun compounds such as *tomato soup* as the combinatory stimuli; given that prior research on single word specificity has focused on nouns, this allowed for a straightforward manipulation of specificity in both positions. If LATL activity primarily reflects the conceptual specificity of the currently accessed item, then the main predictor of LATL amplitudes should be the conceptual specificity of the currently processed item: in the modifier position, activity should be driven by the specificity of the modifier while in the head noun position, the specificity of the head should matter the most. In contrast, if the LATL is primarily a builder of complex meanings, then the specificity of the first word should not matter until the second word is encountered, at which point the meaning of the first word can be integrated with the second. In other words, integrating the modifier *tomato* to a head noun *soup* should elicit a higher LATL amplitude than integrating the modifier *vegetable* to *soup*, given that *tomato* narrows down the meaning of *soup* more than *vegetable* does. Further, on the basis of Westerlund & Pylkkänen's (2014) findings, we would expect conceptually less specific head nouns, such as *dish*, to show larger LATL increases as a function of modifier addition. Overall, under the composition hypothesis, the conceptual specificity of the currently processed item should not by itself affect LATL amplitudes in any substantial way, as already observed in Westerlund & Pylkkänen (2014). To summarize, the aim of the present study was to measure the effects of composition and word-by-word conceptual specificity in order to characterize whether the processing profile of the LATL has a predominantly combinatory or access related nature.

Background: Two semantic effects within the LATL

Concept specificity effect

The hypothesis that the LATL plays a key role in the processing of concepts, and especially specific level concepts, arises from a broad literature involving both neuropsychological data on semantic dementia patients and neuroimaging data on healthy participants.

Semantic Dementia (SD) patients suffer from the loss of semantic memory in both the linguistic and the non-linguistic domains due to progressive atrophy in temporal lobes. Consequently, in behavioral studies, they perform below average (i.e., more slowly and/or less accurately) on tasks requiring conceptual knowledge, such as picture naming, picture sorting, word-picture matching, delayed-copy drawing of a picture, recognizing a distinctive smell or sound, etc. (Snowden et al., 1989; Hodges et al., 1992, 1995; Mummery et al., 1999a, 2000; Garrard & Hodges, 2000; Rogers et al., 2004; Adlam et al., 2006; Garrard & Carroll, 2006; Patterson et al., 2006; Gainotti, 2006, 2007, 2012; Rogers & Patterson, 2007; Lambon Ralph & Patterson, 2008; etc.). However, SD patients' phonological and syntactic knowledge can still be largely preserved (Hodges et al., 1992). Moreover, for SD patients, the loss of semantic memory does not affect the processing of all concepts equally: processing concepts with specific features is more challenging than processing concepts with more general meanings (Warrington, 1975; Hodges et al., 1995; Done & Gale, 1997; Rogers et al., 2005; Rogers & Patterson, 2007). For example, SD patients might still be capable of processing domain-general or prototypical concepts, such as *animal* or *bird*, but unable to process more specific animal

labels (Hodges et al., 1995; Patterson et al., 2006; Rogers & Patterson, 2007). The experiments of Rogers & Patterson (2007) on SD patients even showed that basic-level concepts (e.g., *dog*) turned out to be more challenging than domain-general concepts (e.g., *animal*).

Radiological and PET (positron emission tomography) data have shown the temporal lobes, especially on the left, to be the most consistent location of atrophy in SD patients (Hodges et al., 1992). As regards more detailed localization, Mummery et al. (2000) used VBM (voxel-based morphometry) to show that the most affected brain region of SD patients is the left temporal pole (Brodmann area 38). Further, Mummery et al. (1999a,b) and Rogers et al. (2006) used PET to show that when healthy subjects processed concepts at a specific level, the brain regions activated were consistent with the regions detected in the studies on SD patients, namely the bilateral antero-lateral temporal cortices. Lambon Ralph et al. (2009)'s repetitive transcranial magnetic stimulation (rTMS) study and Binney et al. (2010)'s study across 3 methods (semantic dementia, rTMS and fMRI) provided further evidence confirming that anterior temporal lobes (especially temporal poles) form a critical substrate in the processing of conceptual knowledge.

Behavioral studies on healthy participants have provided further evidence that the specificity level of concepts affects participants' performance in processing concepts. Healthy participants have been reported as faster and more accurate in categorizing concepts at a basic level (e.g., *bird*) than at a more specific level (e.g., *ostrich*) (Rosch et al., 1976; Murphy & Brownell, 1985; Rogers et al., 2006; Rogers & Patterson, 2007). Even though, due to the so-called basic-level advantage (i.e., faster activation of basic-level concepts), healthy participants were in general faster and more accurate in naming or categorizing objects at the basic level than at a more general level, when participants were encouraged to make rapid categorization responses, the basic-level advantage was removed and people were more accurate for domain-general concepts (e.g., *animal*) than basic-level concepts (e.g., *bird*) (Rogers & Patterson, 2007). Brain imaging studies on healthy participants (Mummery et al., 1999a, 2000; Rogers et al., 2006) also confirmed that the LATL is the brain area activated in processing specific concepts, corresponding to the atrophy site of SD patients.

More specifically, within the LATL, the left temporal pole (i.e., left BA38) has been the most robustly attested region for the processing of lexicosemantics and other conceptual information (Mummery et al., 1999a,b; Olson et al., 2007; Desgranges et al., 2007; Lambon Ralph et al., 2009; Binney et al., 2010). Even the study of monkeys' neural activity in processing species-specific calls (Poremba et al., 2004) suggested that significantly greater metabolic activity only occurred in the monkeys' left temporal pole in response to monkey calls.

Other related studies focusing specifically on the role of the left temporal pole in processing linguistic stimuli have shown that the left temporal pole also plays a role in processing words for unique entities and proper names, i.e., concepts of very high specificity (Miceli et al., 2000; Grabowski et al., 2001, 2003; Tranel, 2006, 2009).

Finally, recent fMRI studies focusing on the LATL (Baron et al., 2010; Baron & Osherson, 2011) have shown that the LATL plays a role in processing concept specificity both within words and in word combinations. These two fMRI experiments showed that both the processing of *young man* and the processing of *boy* engaged the LATL more strongly than processing *male*, which suggests that even the processing of basic-level concepts such as *boy* involves more LATL activity than the processing of constituent concepts such as *male*, which are less specific.

In sum, a wide literature has implicated the LATL, and especially the left temporal pole, for conceptual processing at the single concept level, with particular sensitivity to the specificity level of the concept.

Composition effect

In addition to the sizeable literature linking the LATL to conceptual processing at the single concept/word level, the LATL has also been robustly implicated for the combinatory aspects of language processing.

Spatially these effects have shown some variability, but typically they include some or all of left BA38, left BA20, and left BA21.

Specifically, studies contrasting well-formed meaningful sentences and length-matched lists of words or meaningless sentences have consistently shown increased LATL activity for well-formed meaningful sentences, both with hemodynamic methods (Mazoyer et al., 1993; Stowe et al., 1998; Humphries et al., 2001, 2005, 2006; Vandenberghe et al., 2002; Friederici & Kotz, 2003; Rogalsky & Hickok, 2009) and in MEG (Brennan & Pylkkänen, 2012). Moreover, recent MEG experiments have shown that the LATL is also implicated in processing linguistic structures much smaller and simpler than sentences, such as minimal two-word phrases (Bemis & Pylkkänen, 2011, 2013a,b; Westerlund & Pylkkänen, 2014; Westerlund et al., 2015). For example, processing *blue boat* (a meaningful well-formed combinatorial phrase) caused more activity in the LATL than processing *xptq boat* (a non-combinatorial length-matched one-word condition) or *cup boat* (a non-combinatorial length-matched word list condition) (Bemis & Pylkkänen, 2011). In these MEG experiments, the composition effect was attested in the time window about 200 ms to 300 ms after the onset of the target word (in this example, *boat*).

A central question within the combinatory research on the LATL is whether it reflects syntactic or semantic-aspects on composition. A popular paradigm for addressing this has been the so-called jabberwocky version of the sentence vs. list manipulation, where all the open class words of the stimuli are replaced by pseudowords, with the intent of removing semantics from the expressions. While in several studies, jabberwocky sentences have elicited a LATL increase (Friederici et al., 2000; Humphries et al., 2006; Mazoyer et al., 1993), arguably suggesting a more syntactic role (for discussion see Del Prato & Pylkkänen, 2014), this has not always been the case (Pallier et al., 2011). A syntactic explanation is also challenged by the results of Vandenberghe et al. (2002), who contrasted grammatical and scrambled versions of semantically well-formed and anomalous sentences, showing that the presence of syntactic structure is not sufficient for a LATL increase, rather, the sentence also needs to be semantically meaningful. A similar conclusion can be drawn from the already discussed Westerlund & Pylkkänen (2014) results where a combinatory LATL increase was observed only for conceptual less specific nouns; again a syntactic account would have predicted a LATL increase whenever composition was present. Finally, Del Prato and Pylkkänen (2014) have shown that while adjectival modification increases LATL amplitudes, numeral quantification does not (e.g., *red cups* vs. *two cups*). Since both types of phrases involve syntactic composition, this difference also must have a semantic as opposed to a syntactic origin. In all, current evidence heavily points towards a semantic combinatory role for the LATL.

In light of the evidence that the LATL's contribution to composition is likely semantic as opposed to syntactic, the relatively early timing of its composition effects at about 200-300 ms is somewhat surprising. Presumably, semantic composition cannot occur before lexical-semantic access has taken place and while some electrophysiological evidence does indicate lexical-semantic factors to come into play before 200 ms (e.g., Pulvermüller et al., 2001, 2005; Shtyrov et al., 2004; Pulvermüller, 2005), the most consistent time window for such effects is later, at 300–500 ms, i.e., in the classic "N400 time window" (e.g., Kutas & Van Petten, 1988, 1994; Kutas & Federmeier, 2000, 2009, 2011; Pylkkänen & Marantz, 2003; Van Petten & Luka, 2006; Federmeier, 2007; Lau et al., 2008, 2009). Together, though, these two bodies of lexical-semantic results raise the possibility that lexical-semantic activation may in fact unfold gradually over time (cf., Binder and Desai, 2009), as opposed to in a single processing stage. This could subsequently entail that semantic composition may also occur in several stages, differing in the depth of semantic processing that has occurred on the input items. In this type of hypothesis, the 200–300 ms composition effects in the LATL could index a relatively early combinatory stage, potentially followed by later ones (Pylkkänen, in press). Thus understanding the level of semantic analysis in early LATL activity is clearly

an important topic for future studies, but one that the current work did not directly aim to tackle.

Current study: The temporal dynamics of specificity effects word by word

Given that multiword utterances often have more specific meanings than single words or other noncombinatory stimuli, it is conceivable that the LATL effects of composition and single word conceptual specificity both relate to a unified mechanism of meaning specification (Westerlund & Pylkkänen, 2014). In the MEG experiments of basic combinatorial phrases (Bemis & Pylkkänen, 2011, 2013a,b), combining the word *boat* with an adjectival modifier *blue*, a more specific concept – *blue boat* – is generated: in this more specific concept, the modifier – *blue* – highlights the color dimension of the object – *boat* – and specifies the value of this dimension (Murphy, 2002). The studies of Baron et al. (2010) and Baron & Osherson (2011) could be interpreted in a similar way. These two studies investigated conceptual combination both across-words (e.g., *young man* vs. *man*) and within words (e.g., *boy* vs. *male*): by specifying the value in the dimension of age, the resultant concepts (*young man* and *boy*) are more specific than *male* with regard to this dimension.

Interpreting the LATL combinatory effect as an instance of the conceptual specificity effect conforms to the "semantic hub" model of the LATL (Patterson et al., 2007), where attributes of a concept (such as name, action, color, shape, motion, etc.) are distributed over a wide neural network within which the LATL acts as an organizing "hub", connecting the various attributes into a holistic concept. Therefore, for the semantic hub, processing a more specific concept means recruiting more features and evaluating the values of these features, and thus, this model explains why processing a more specific concept is more costly for the semantic hub than processing a more general concept. And it follows directly why the left temporal pole is also implicated in processing concepts of unique entities or proper names (Miceli et al., 2000; Grabowski et al., 2001, 2003): in processing these highly specific concepts, more features are recruited and values have to be evaluated with a high precision.

To investigate whether the specificity and composition effects co-localize and/or interact, Westerlund & Pylkkänen (2014) compared the MEG activity elicited by processing concepts of higher (e.g., *trout*) vs. lower specificity (e.g., *fish*) in both a combinatorial context (e.g., *spotted fish / trout*) and a non-combinatorial context (e.g., *xhsl fish / trout*). As already summarized, the results indicated that even though noun specificity (i.e., *trout* vs. *fish*) only subtly modulated LATL activity in the non-combinatorial context, it robustly affected the size of the composition effect in the combinatorial context, with low specificity nouns eliciting a larger effect (i.e., the difference between *spotted fish* and *xhsl fish* is robustly larger than the difference between *spotted trout* and *xhsl trout*). Post-hoc Amazon Mechanical Turk norming tests ruled out some confounding factors, such as imageability and plausibility of combinatorial phrases. Therefore, the results suggested that the adjectival modifier plays a more measurable role in specifying the meaning of a low specificity noun than in specifying the meaning of a noun that is already relatively high in specificity. Thus the increase of brain activity in the LATL may be driven by the change in specificity, i.e., the specificity difference between a low specificity concept in isolation and in a combinatorial context (here *fish* and *spotted fish*) is larger than the specificity difference between a high specificity concept in isolation and in a combinatorial context (here *trout* and *spotted trout*).

In order to further investigate the interplay between the effects of concept specificity and composition on the LATL, the current study employed a design where conceptual specificity was varied in both the modifier and head noun position, to achieve a fuller characterization of the temporal dynamics of the two factors. Specifically, we employed a 2 by 3 design on noun-noun combinations, with the specificity of the modifier (consonant string vs. general vs. specific, e.g., *xpt* vs. *vegetable*

vs. *tomato*) and the specificity of the head word (general vs. specific, e.g., *dish* vs. *soup*) as main factors (Table 1). Thus the study consisted of four combinatorial conditions and two single-word (baseline) conditions. In light of the prior finding that the conceptual specificity of the head noun affects the composition effect (Westerlund & Pylkkänen, 2014), we specifically aimed to characterize the effect of modifier specificity on the combinatory response. If the size of the combinatory effect indeed reflects the proportional change in specificity induced by the addition of a modifier, then the largest combinatory effect should be elicited for less specific head nouns in the context of a more specific modifier. In contrast, if LATL amplitudes primarily reflect the conceptual specificity of the currently processed item, modifier specificity might impact LATL amplitudes primarily at the modifier but not necessarily at its combinatory site, i.e., the head noun. With the millisecond temporal resolution of MEG we were able to characterize the impact of our specificity manipulation throughout the two-word phrase.

Materials and methods

Participants

23 right-handed, native English speakers participated in the study. All had normal or corrected-to-normal vision and gave informed consent. One participant was excluded from MEG data analysis due to marker coils falling off during the recording; a second participant was excluded due to excessive noise during the MEG recording; a third participant was excluded due to the malfunction of the trigger box; and a fourth participant was excluded because more than 90% of the trials had to be rejected in data processing due to blinking. Thus 19 participants were included in the final analysis (10 female; average age = 24.95 years, SD = 5.21 years). All data were collected at the Neuroscience of Language Lab at New York University in New York.

Experimental design and stimuli

The experiment employed a 2 by 3 design, with 60 groups of target words forming a total of 360 trials. However, only 48 of these groups were included in the analysis as a more detailed statistical analysis of the stimulus properties revealed that there were 12 sets that created confounds to the design. Each group contained a pair of modifiers (general vs. specific, e.g., *vegetable* vs. *tomato*) and a pair of head words (general vs. specific, e.g., *dish* vs. *soup*). As Table 1 shows, these four word types formed the stimuli of the six conditions employed in the experiment. In order to match the amount of visual stimulation prior to the head word in all the 6 conditions, we inserted unpronounceable consonant strings before the head noun in those conditions where there was no modifier, following the procedure in prior studies using this paradigm (e.g., Bemis & Pylkkänen, 2011; Westerlund & Pylkkänen, 2014). Therefore, in the whole experiment, each modifier appeared twice, in two trials respectively; and each head word appeared three times, in three trials respectively. With such a design, we gained a good control of the specificity factor between the conditions: for example, the difference between the GenGen condition and SpecGen condition was similar to the difference between the GenSpec condition and SpecSpec condition in that, in both cases, we were using the same Gen modifier and the same Spec modifier and the same head words. In our experiment,

in each pair of words, the general meaning word might be a domain-general word or a basic-level word, and the specific meaning word might be a basic-level word or a subordinate level word. In other words, the specificity of a word was defined in a relative way: in each pair, the word was defined as general or specific when it was compared to the other word in the pair. We used Wordnet Search 3.1 (http://wordnet.princeton.edu/) (Miller et al., 1990; Miller, 1995), a large lexical database showing subordinates and superordinates for each English word, to check the specificity relation of all the pairs of words used in the experiment.

Words of the critical stimuli were matched for the following six factors (see Table 2):

(1) length (values from the English Lexicon Project (Balota et al., 2007));
(2) number of syllables (values from the English Lexicon Project);
(3) number of morphemes (values from the English Lexicon Project);
(4) lexical decision reaction time (values from the English Lexicon Project);
(5) naming reaction time (values from the English Lexicon Project);
(6) log frequency (calculated from the data in the Corpus of Contemporary American English (Davies, 2008) by following Balota et al. (2007)).

Furthermore, the 4 two-word combinatorial conditions were also matched for the following two factors (see Table 3):

(1) bigram frequency (values from the Corpus of Contemporary American English (COCA) (Davies, 2008));
(2) transition probability (calculated from the COCA data).

Fig. 1 illustrates the trial structure. Each trial contained a fixation "+", a modifier (or a meaningless consonant string), and a head word. Each of these items lasted 300 ms on the screen, and the interval between two neighboring items was also 300 ms. In order to keep participants active, prevent them from feeling sleepy and verify that they were indeed reading and processing the words appeared on the screen, we added a Yes/No question to one third of the 360 trials. Each question was not longer than 4 words, presented with capitals and followed by a question marker: the question could be related to the modifier or the head word only, or to the phrase as a whole. The 120 questions were evenly distributed among the 6 conditions; the correct answer to half of the questions should be Yes and the correct answers to the other half should be No; the Yes and No were also evenly distributed among the 6 conditions. The question stayed on the screen until the participant pressed a button to answer it. Participants were asked to try not to blink when reading the words shown on the screen. Thus, in order to give participants more opportunities to blink, for every three trials, we added one BLINK section between the fixation and the 1st word position; the BLINK section also lasted 300 ms, and the interval between the fixation and the BLINK section and the interval between the BLINK section and the 1st word position were 300 ms each.

The stimuli were presented by PsychToolBox software (Brainard, 1997; Pelli, 1997). The target items were presented in lowercase letters;

Table 1
Experimental design and illustration of the stimuli.

	Consonant string (xpt)	General meaning modifier (GenMod)	Specific meaning modifier (SpecMod)
General meaning head word (GenHead)	xptGen: *qptg dish*	GenGen: *vegetable dish*	SpecGen: *tomato dish*
Specific meaning head word (SpecHead)	xptSpec: *qptg soup*	GenSpec: *vegetable soup*	SpecSpec: *tomato soup*

Table 2

Statistics of critical stimuli (single words). The words are matched for length, number of syllables, number of morphemes, lexical decision reaction time, naming reaction time, and COCA log frequency: the means, standard deviations and the p values of 2 by 2 ANOVAs are shown in this table – with the position (modifier vs. head word) and the specificity (low/Gen vs. high/Spec) as main factors. No significant differences are observed among the stimuli for these factors.

	Num	Length mean (SD)	NSyll mean (SD)	NMorph mean (SD)	LD RT mean (SD)	NMG RT mean (SD)	COCA log freq mean (SD)
Gen mod	48	5.73 (1.71)	1.75 (0.81)	1.21 (0.46)	635.55 (68.44)	628.03 (51.45)	9.14 (1.30)
Spec mod	48	5.71 (1.89)	1.85 (0.77)	1.17 (0.38)	645.11 (76.08)	631.80 (58.96)	8.85 (1.47)
Gen head	48	6.21 (1.99)	1.94 (0.89)	1.35 (0.56)	638.02 (71.57)	638.29 (59.57)	9.25 (1.46)
Spec head	48	5.54 (1.69)	1.63 (0.70)	1.25 (0.53)	642.47 (81.55)	635.01 (59.30)	8.92 (1.83)
ANOVA p (position)		0.5535	0.8663	0.1047	0.9692	0.4110	0.6851
ANOVA p (specificity)		0.1932	0.3656	0.3008	0.4807	0.9765	0.1697
ANOVA p (interaction)		0.2214	0.0713	0.6571	0.7738	0.6668	0.9259

the Yes/No questions and the BLINK were presented in capital letters; all of the characters were presented in white 30-point Courier font on a grey background.

The 360 trials were divided into 12 blocks: each block contained 30 trials (5 trials for each condition); any modifier or head word appeared no more than once in one block; among the 30 trials of one block, roughly 9 to 12 of the trials contained a Yes/No question; between two neighboring trials, the interval was normally distributed with a mean of 400 ms (SD = 100 ms). The trials within each block were randomized and the blocks were also randomized in the experiment for each participant. However, the BLINK section was inserted regularly: it appeared in the 2nd, 5th, 8th… 29th trials in each block. Between blocks, participants could choose to take a short rest or continue immediately.

Procedure

Before the experiment, we used a Polhemus Fastscan three-dimensional laser digitizer to scan participants' head shapes, 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 five coils located across the forehead with respect to the MEG sensors.

Then, before the MEG recording, participants practiced the task on a shortened block (including 12 trials, 5 of which had a Yes/No question) outside the magnetically-shielded room (MSR). During this practice session, participants were given the same instructions as in the real experiment: they were told to press a button with either the index finger or the middle finger of their left hand to answer the question. Participants got feedback after they answered each question in the practice session, so that participants could verify their comprehension of the task and familiarize themselves with the experiment process. The practice session could be repeated when participants wanted to get more practice.

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 157-channel axial gradiometer system (Kanazawa Institute of Technology, Nonoichi, Japan), at a 1000Hz sampling rate with a low-pass filter at 200Hz and a notch filter at 60Hz. Stimuli were projected onto a screen about 50 cm away from participants' eyes. As in the practice session, during the recording, participants used the index finger and the middle finger of their left hand to answer questions, but they got no feedback. The recording session lasted approximately 20 minutes.

Data preprocessing

In MEG160 (Meg Laboratory 2.004A), raw MEG data were first noise-reduced and filtered with a high-pass filter of 1Hz. Then MEG data from the 200 ms before the onset of each modifier (in two-word conditions) or the onset of each consonant string (in one-word conditions) (i.e., 200 ms pre-stimulus (or −200 ms)) to 1300 ms after this onset (i.e., 1300 ms post-stimulus (=700 ms after the onset of the head word)) were segmented out. In MEG160, raw data of each segment were also inspected for blinks and data segments containing blinks were manually rejected. After epoching, the digitized head shape information and marker coils' locations were used to constrain the source localization by co-registering the five coils located around the face with respect to the MEG sensors. In BESA 5.1 (MEGIS Software GmbH, Gräfelfing, Germany), we further rejected trials that exceeded an amplitude threshold of 3000fT and generated average files per condition per participant. A low-pass filter at 40Hz was applied to averages.

We separately created averages for the 1st-position words (i.e., modifiers) and for the 2nd-position words (i.e., head words). For each position, the baseline was defined from 100 ms pre-onset to the onset (i.e., −100 ms to 0) and the artifact rejection was defined from −100 ms to 600 ms.

Minimum Norm Estimates

After pre-processing, distributed L2 minimum norm source estimates were constructed for each average using BESA 5.1 (MEGIS Software GmbH, Munich, Germany). The channel noise covariance matrix for each estimate was based upon the 100 ms prior to the onset of the noun in each condition average. The construction of the source estimates was based on the activity of 1426 regional sources evenly distributed on two shells (713 regional sources each shell), at 10 and 30% below a smoothed standard brain surface adjusted to the head shape information gathered during the laser digitization. Regional sources in MEG can be thought of as two single dipoles at the same location but with orthogonal orientations. The resultant activity at each regional source was computed as the root mean square of the source activities of its two components. After this, pairs of regional sources in the outer and inner shells of the source space were compared and the source with the larger value was selected for further computation. Thus overall, the source space consisted of 713 non-directional sources. Minimum norm images were depth as well as spatiotemporally weighed, using a signal subspace correlation measure (Mosher & Leahy 1998).
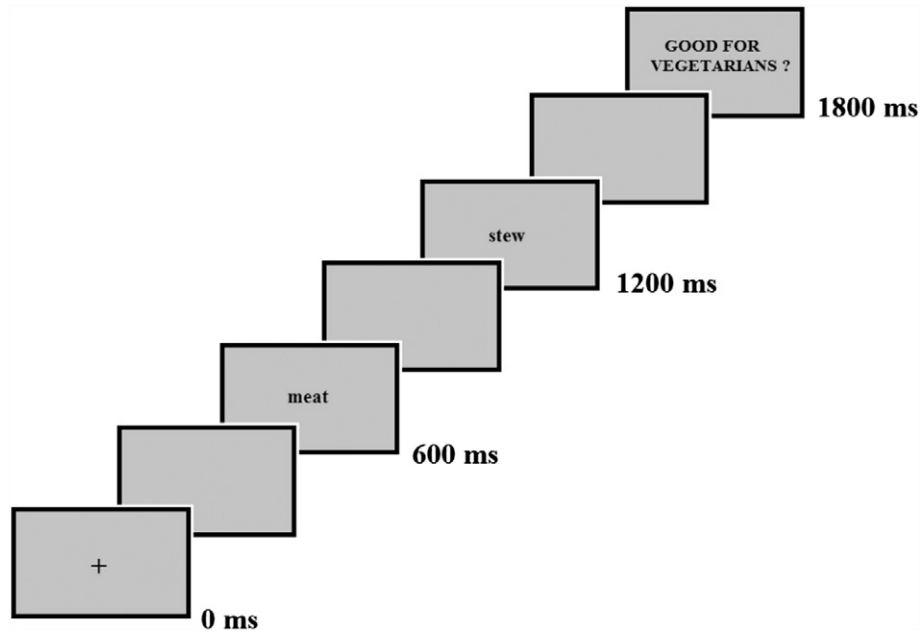
Table 3

Statistics of critical stimuli (two-word phrases). The phrases are matched for COCA bigram frequency and COCA transition probability: the means, standard deviations and the p values of 2 by 2 ANOVAs are shown in this table – with modifier specificity and head word specificity as main factors. No significant differences are observed among the stimuli for these factors.

	Num	COCA bigram frequency mean (SD)	COCA transition probability mean (SD)
GenGen	48	36.25 (58.39)	0.003805 (0.005928)
SpecGen	48	27.38 (79.99)	0.003777 (0.010023)
GenSpec	48	23.5 (49.62)	0.004572 (0.011005)
SpecSpec	48	37.85 (37.85)	0.005459 (0.014463)
ANOVA p (modifier specificity)		0.7926	0.7829
ANOVA p (head specificity)		0.9132	0.4329
ANOVA p (interaction)		0.2657	0.7692

Fig. 1. Trial structure.

Data analysis

Our analyses were focused on the following two measures:

(1) the LATL activity generated by the processing of the 2nd-position words (i.e., head words) in these six conditions: xptGen_2nd, GenGen_2nd, SpecGen_2nd, xptSpec_2nd, GenSpec_2nd, SpecSpec_2nd. These measurements allowed us to perform ANOVAs to test the interaction between the main factor of head specificity and the main factor of modifier specificity.
(2) the LATL activity generated by the processing of single noun modifiers: GenGen_1st, SpecGen_1st, GenSpec_1st, SpecSpec_1st. In these measurements, the target words are the 1st-position words (i.e., modifiers). Together with the measurements of xptGen_2nd and xptSpec_2nd, these measurements allowed us to perform *t*-tests to test the single-word specificity for the head word position and the modifier position respectively.

First we performed ROI (region-of-interest) analyses focusing on the left temporal pole on a specific time window, so as to determine the spatiotemporal dynamics of processing single words and combinatorial phrases. This was followed by whole-brain source analyses across all sources and time points to verify that the effects observed in the ROI analyses in fact reflected activity within left BA38 as opposed to neighboring regions.

ROI analysis

Based on previous studies on the concept specificity effect (see Section 1.1.1), our ROI analysis focused on left BA 38 (i.e., the left temporal pole), though given the spatial resolution of MEG (which can vary from a few millimeters in optimal circumstances to very poor in the case of e.g. deep sources, Hämäläinen et al., 1993), it should be kept in mind that this ROI could also reflect activity from neighboring regions (but see the description of our full brain contrasts below as a way to partly address this). To isolate BA 38 within our source space, the Tailarach daemon (Lancaster et al., 1997, 2000) was used to partition the 713 sources on the smooth BESA cortex into Brodmann area labels. All sources within the left BA38 were first averaged together and then

the time course of this ROI was submitted to nonparametric, cluster-based permutation tests (Maris & Oostenveld, 2007) aimed at identifying temporal clusters of activity where our stimulus manipulation reliably affected the source intensity, corrected for multiple comparisons. The permutation tests were ran over a 150 – 350 ms time window for each target word, motivated by the timing of LATL composition effects in prior studies, typically observed at 200–250 ms post-stimulus (Bemis & Pylkkänen, 2011, 2013a,b; Westerlund & Pylkkänen, 2014).

The basic idea of the permutation testing is to calculate test statistics on randomly partitioned data many times (here 10,000 times) and then to calculate the number of random partitions that yield a test statistic at least as high as the observed test statistic. If the number of such partitions is smaller than the pre-set *p*-value (usually 0.05), then the effect is treated as "significant".

The specific steps in our cluster-based permutation tests are as follows:

(1) for every time point, an uncorrected *t*-test or ANOVA was calcualted on the ROI time course data;
(2) within our time window (150 to 350 ms for each target word), clusters of at least 10 adjacent time points that showed a significant effect at a *p*-value of 0.05 (uncorrected) were selected;
(3) a test statistic (for *t*-tests, it is the sum of *t*-values of every time point in a cluster; for ANOVAs, it is the sum of F-values of every time point in a cluster) was calculated for each cluster of time-points passing the criteria defined in step (2) above and the cluster with the largest test statistic within the analyzed time-interval (i.e., 150–350 ms post word-onset) was selected;
(4) then the data were permuted for 10,000 times by randomly re-assigning the condition labels for each participant independently, and the test statistic was calculated for each largest cluster in this permutation; thus, these 10,000 test statistics formed the null distribution for the test statistic of interest;
(5) if less than 5% of the partitions yield a test statistic larger than the observed test statistic of (3), than the cluster was accepted as significant.

To investigate the main factor of head specificity and the main factor of modifier specificity, we first ran a 2 by 3 permutation ANOVA over all

the six conditions over the time window 150 ms to 350 ms at the head word presentation (i.e., focusing on the 2nd position) in left BA38 with the specificity of the head word and the specificity of the modifier as two main factors.

Since the 2 by 3 ANOVA revealed that there was a main effect of the modifier specificity, as a follow-up analysis, we also conducted a 2 by 2 permutation ANOVA over the four combinatorial conditions over the same time window (i.e., 150 ms to 350 ms at the head word presentation), also with the specificity of the head word and the specificity of the modifier as two main factors. The purpose of this follow-up 2 by 2 ANOVA was to confirm that the modifier specificity effect shown in the 2 by 3 ANOVA was not simply due to the existence of modifiers in two-word conditions, but really due to the difference in the modifier specificity.

Another follow-up analysis was to verify the composition effect, contrasting combinatory conditions to single word conditions. Specifically, four permutation t-tests were performed over the time window 150 ms to 350 ms at the head word presentation in left BA38 – xptGen vs. GenGen, xptGen vs. SpecGen, xptSpec vs. GenSpec, xptSpec vs. SpecSpec.

Finally, two permutation t-tests were performed to investigate the single-word specificity of different positions: (1) the comparison between general modifiers and specific modifiers over the time window 150 – 350 ms at the modifier presentation; (2) the comparison between the condition xptSpec and the condition xptGen over the time window 150 – 350 ms at the head word presentation.

Since the permutation t-test between xptSpec and xptGen showed a subtle effect for the single-word specificity effect, to further examine the size of this effect, a follow-up t-test was performed on averaged activity over the time window of this cluster (i.e., 266 - 275 ms) at the head word presentation.

Whole brain analysis

In order to reveal any robust effects outside of our ROI and to assess whether the effects shown in the ROI analyses in fact reflected activity in the left temporal pole as opposed to spill-over activity from adjacent regions, we performed four whole brain analyses across the entire source space, as follows:

(1) in order to verify the main effect of modifier specificity, we performed a comparison between conditions containing a specific modifier and conditions containing a general modifier for the time window of 0 to 600 ms after the head word presentation (i.e., the 2nd position);

(2) to verify the significant composition effect shown in the permutation t-test between the two conditions xptGen and SpecGen, we performed a comparison between these two conditions for the time window of 0 to 600 ms at the head word presentation (i.e., the 2nd position).

(3) to verify the results of ROI analyses on single-word specificity, we performed a comparison between all specific modifiers and all general modifiers for the time window of 0 to 600 ms at the modifier presentation (i.e., the 1st position);

(4) finally we performed a comparison between xptSpec and xptGen for the time window of 0 to 600 ms at the head word presentation (i.e., the 2nd-position).

In each of these full brain analyses, a paired two-tailed t-test was conducted sample by sample for every time point over the targeted time window. Results were plotted as significant when at least 5 adjacent sources and at least 5 consecutive time points (5 ms) showed a difference at a p-value of 0.05 (uncorrected for multiple comparisons). As emphasized above, the purpose of these analyses was simply confirmatory of the ROI analyses.

Sensor space analyses

In order to assess whether our source results are also observable in the sensor data, we divided the left hemisphere sensors into four even sections (like a pie), moving from anterior to more posterior sensors, and calculated the root mean square (RMS) for each subsection. The sensor division was intended as maximally mechanical and thus easily replicable, though the downside of this was a loss of sensitivity, as the analysis was not designed to capture variability in individual subjects' magnetic field patterns.

RMSed data were analyzed with permutation tests as the BA 38 source data, although given the noisier nature of the sensor data (with each sensor likely reflecting activity from multiple sources) we increased the p-value used for cluster selection to .3 and conducted the permutation tests within the narrow time window of 200 – 300 ms, i.e., directly focused around the LATL effects observed in the source analysis. For simplicity, all tests were one-tailed t-tests, directly aimed at assessing either effects of composition (i.e., increases for combinatory conditions as compared to their one-word controls) or increases for specific over general words in non-combinatory context (i.e., specific vs. general modifiers in the 1st word position and specific vs. general single words in 2nd word position).
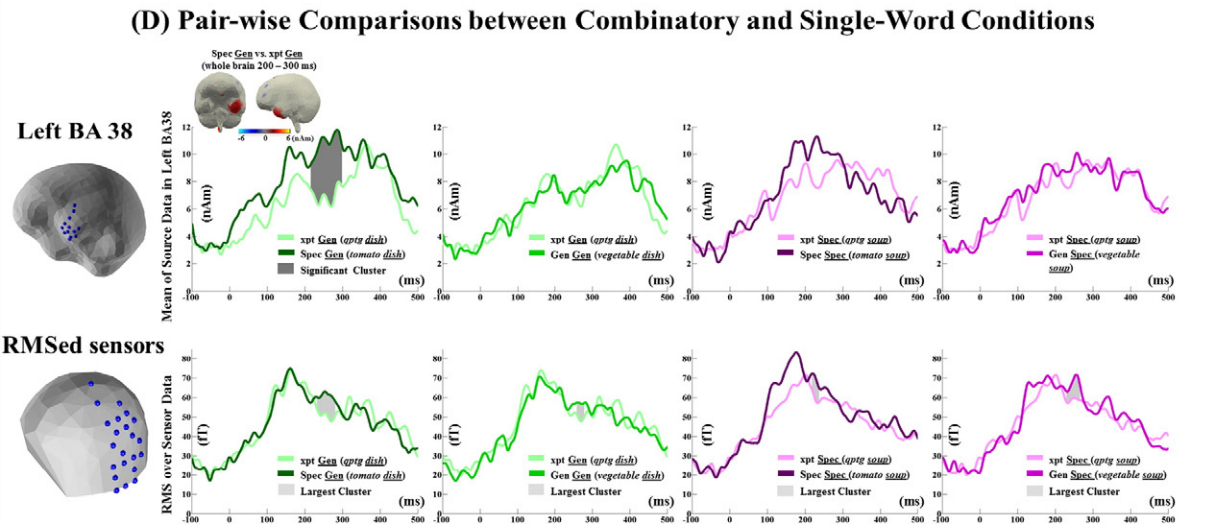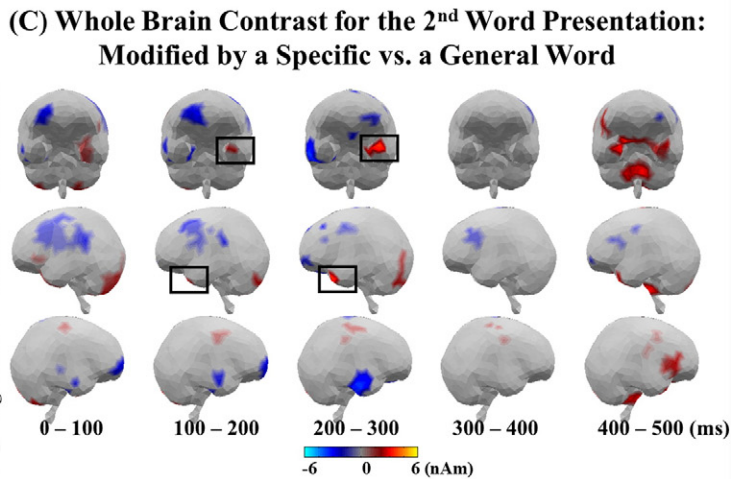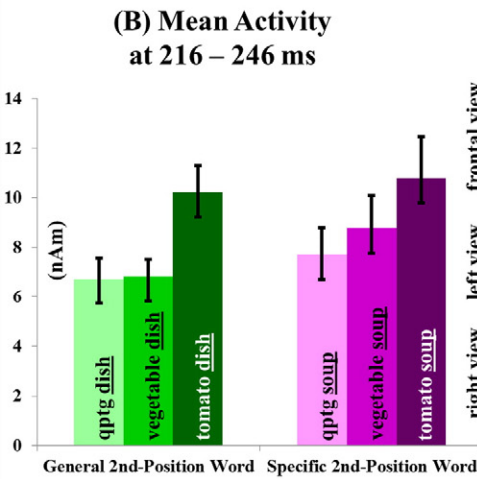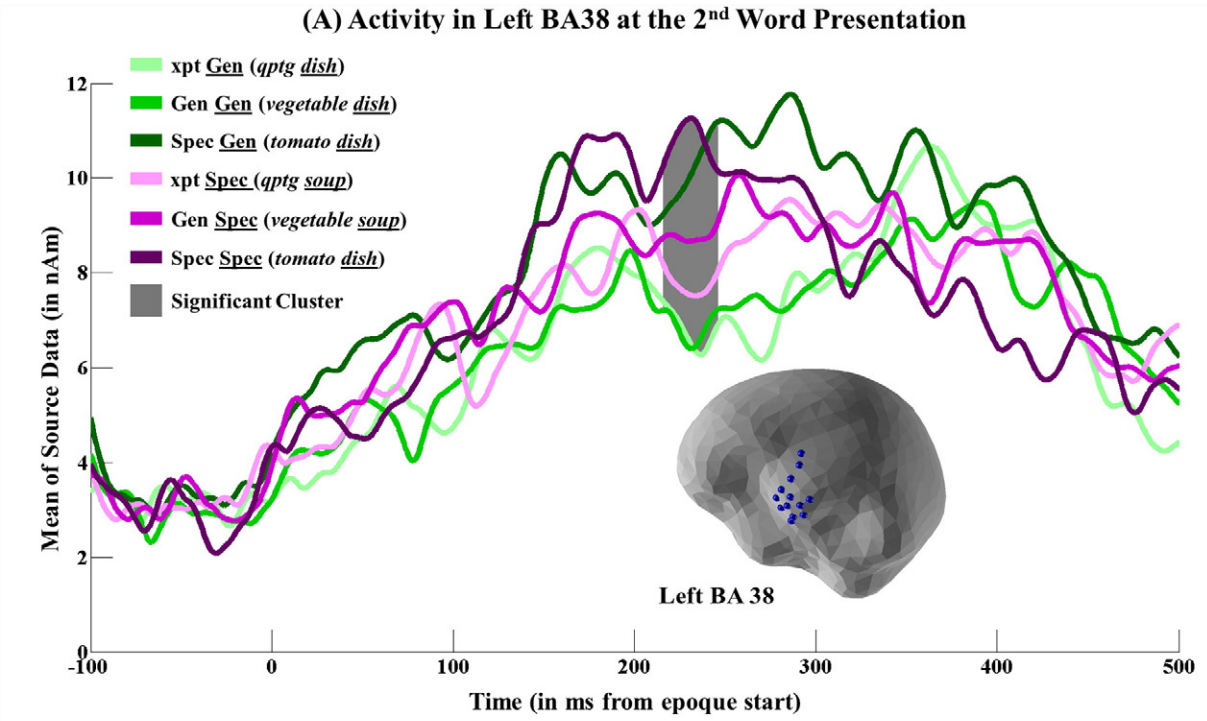
Results

Behavioral results

Of the total 360 trials (including 72 fillers and 288 critical trials), 1/3 included a task – a Yes/No question, so there were 120 questions. Mean accuracy for all 19 participants was 88.07% ± 2.3 (standard deviation). Since in this experiment, the purpose of including the task was simply to keep participants paying attention and prevent them from feeling bored and sleepy, no further analysis was done for the behavioral results.

ROI Results

The interplay between the concept specificity effect and the composition effect

The 2 by 3 permutation ANOVA performed over all the six conditions (i.e., xptGen, GenGen, SpecGen, xptSpec, GenSpec, SpecSpec) on the left BA38 in the time window 150 – 350 ms at the head word presentation (i.e., the 2nd position) revealed a significant cluster between 216 ms and 246 ms ($p = 0.037$) for the main effect of the modifier specificity (Fig. 2). Within this cluster, increasing the specificity of the modifier led to an increase of brain activity in left BA38 (xptGen Mean = 6.735 (SE = 0.841), GenGen Mean = 6.826 (SE = 0.694), SpecGen Mean = 10.215 (SE = 1.071), xptSpec Mean = 7.712 (SE = 1.078), GenSpec

Fig. 2. The impact of modifier specificity on the composition effect at the head word. (A) Activation (in nAm) by condition in the left BA38, where 0 ms represents the onset of the 2nd position word. A 3D BESA standard brain illustrates the sources included in the ROI, with sources color-coded by Brodmann area. A significant cluster for the main effect of the modifier specificity was found at 216 – 246 ms ($p = 0.037$). (B) Average activity over the interval during which the reliable main effect of modifier specificity is observed (error bars show SEMS). (C) Uncorrected whole brain comparisons at the 2nd word presentation contrasting conditions modified by a more specific word vs. by a less specific word. Activity is plotted in red if the more specific modifier elicited greater activity than the less specific one and in blue if the reverse was true. (D) Pairwise comparisons between combinatory conditions and their single-word control conditions. The top panel shows the comparisons of source data in left BA38 (in nAm). A significant composition cluster (i.e. increased activity for the two-word condition) was only found for the SpecGen vs. xptGen comparison (216 – 298 ms, $p = 0.004$). The bottom panel shows the comparisons of sensor data (in fT) (a 3D BESA helmet illustrates the sensors included in the analyses). Similarly to the BA 38 source data, a cluster of increased activity is observed between 200 ms and 300 ms for the SpecGen combinatory condition as compared to its one-word control (xptGen), though statistically this effect was weaker than the source effect (see Sections 2.6.3 and 3.4).

(A) Activity in Left BA38 at the 2nd Word Presentation

Legend:
- xpt Gen (*qptg dish*)
- Gen Gen (*vegetable dish*)
- Spec Gen (*tomato dish*)
- xpt Spec (*qptg soup*)
- Gen Spec (*vegetable soup*)
- Spec Spec (*tomato dish*)
- Significant Cluster

Left BA 38

Mean of Source Data (in nAm) vs. Time (in ms from epoque start)

(B) Mean Activity at 216 – 246 ms

(nAm); General 2nd-Position Word: qptg dish, vegetable dish, tomato dish; Specific 2nd-Position Word: qptg soup, vegetable soup, tomato soup

(C) Whole Brain Contrast for the 2nd Word Presentation: Modified by a Specific vs. a General Word

frontal view / left view / right view

0 – 100   100 – 200   200 – 300   300 – 400   400 – 500 (ms)

-6   0   6 (nAm)

(D) Pair-wise Comparisons between Combinatory and Single-Word Conditions

Left BA 38

Spec Gen vs. xpt Gen (whole brain 200 – 300 ms)
-6   0   6 (nAm)

Mean of Source Data in Left BA38 (nAm)

- xpt Gen (*qptg dish*)
- Spec Gen (*tomato dish*)
- Significant Cluster

- xpt Gen (*qptg dish*)
- Gen Gen (*vegetable dish*)

- xpt Spec (*qptg soup*)
- Spec Spec (*tomato soup*)

- xpt Spec (*qptg soup*)
- Gen Spec (*vegetable soup*)

RMSed sensors

RMS over Sensor Data (fT)

- xpt Gen (*qptg dish*)
- Spec Gen (*tomato dish*)
- Largest Cluster

- xpt Gen (*qptg dish*)
- Gen Gen (*vegetable dish*)
- Largest Cluster

- xpt Spec (*qptg soup*)
- Spec Spec (*tomato soup*)
- Largest Cluster

- xpt Spec (*qptg soup*)
- Gen Spec (*vegetable soup*)
- Largest Cluster

Mean = 8.775 (SE = 1.318), SpecSpec Mean = 10.8 (SE = 1.667)). No clusters showed sensitivity to head word specificity or the interaction between the two main factors.

The follow-up 2 by 2 permutation ANOVA performed over the four two-word conditions (i.e., GenGen, SpecGen, GenSpec, SpecSpec) in the same time window (i.e., 150 – 350 ms at the head word presentation) in left BA38 showed an almost significant cluster for the main effect of the modifier specificity between 218 ms and 244 ms ($p = 0.06$). Also within the cluster, the pattern showed that increasing the modifier specificity caused an increase of brain activity in left BA38. This result confirmed that the main effect of the modifier specificity shown in the 2 by 3 ANOVA was not simply due to the existence of a modifier in two-word conditions, but largely due to the specificity of the modifier.

The results of the four permutation *t*-tests performed between combinatory and single-word conditions at the head word on the left BA38 in the time window of 150 – 350 ms were as follows: only the comparison between xptGen and SpecGen generated a significant cluster between 216 ms and 298 ms ($p = 0.004$, more brain activity was elicited for SpecGen than for xptGen); for the comparisons between xptGen and GenGen, between xptSpec and SpecSpec, and between xptSpec and GenSpec, no clusters were found (Fig. 2). The results of these *t*-tests showed that a significant composition effect was only elicited when the head word was general and the modifier was specific.

The concept specificity effect of single words

The permutation *t*-test performed between all specific modifiers and all general modifiers over the time window 150 ms to 350 ms at the modifier presentation on left BA38 showed no clusters (see Fig. 3).

The permutation *t*-test performed between xptSpec and xptGen in the time window 150 ms to 350 ms at the head word presentation on left BA38 showed a small cluster between 266 ms and 275 ms (which did not however survive correction for multiple comparison; $p = 0.23$, corrected) with a pattern showing that specific head words caused more neural activity than general head words (see Fig. 3). When probed further, the follow-up window *t*-test performed on averaged activity between 266 ms and 275 ms showed a significant effect ($p = 0.03$).

Whole brain results

The interplay between the concept specificity effect and the composition effect

Based on the results of ROI analyses, we performed a whole brain comparison between conditions containing specific modifiers and conditions containing general modifiers over the time window 0 ms to 600 ms at the head word presentation (i.e., the 2nd position). This comparison confirmed the main effect of the modifier specificity in the left temporal pole: the clusters were plotted in red during 200 ms to 300 ms in the left temporal pole, which means that in the left temporal pole, the conditions with specific modifiers led to more activity than condition with general modifiers. During 100 ms to 200 ms, there was also a small cluster in the left temporal pole (Fig. 2). Most other effects observed for this contrast were in the other direction, i.e., showing more activity for general than specific modifiers (plotted in blue).

Another whole brain comparison was performed between the conditions xptGen and SpecGen over the time window 0 ms to 600 ms at the head word presentation (i.e., the 2nd-position), to examine the composition effect. This comparison also confirmed the significant composition effect in the left temporal pole: the clusters were plotted in red during 200 ms to 300 ms in the left temporal pole, showing that in the left temporal pole, adding a specific modifier to a general head word led to more brain activity (see Fig. 2).

The concept specificity effect of single words

Two whole brain comparisons were performed to verify the results of ROI analyses on single-word specificity effect. The comparison

between all specific modifiers and all general modifiers over the time window 0 ms to 600 ms at the modifier presentation (i.e., the 1st-position) showed no cluster in LATL, which means that no single-word specificity was attested for modifiers (see Fig. 3). However, this whole brain comparison did show an increase for specific over general modifiers in several other regions, such as dorsal parts of left motor cortex, the left inferior frontal gyrus and ventromedial prefrontal cortex (see Fig. 3). Thus although increased specificity in the modifier position did not modulate activity in the left anterior temporal lobe, it did have a rather distributed effect in many other regions.

The whole brain comparison between xptSpec and xptGen over the time window 0 ms to 600 ms at the head word presentation (i.e., the 2nd position) showed a very small cluster during 200 ms to 300 ms in the left temporal pole, which conformed to the result of the corresponding ROI analysis (See Fig. 3): single head words of high specificity subtly led to more brain activity in the left temporal pole than single head words of low specificity. No other regions showed any obvious increases. Also, overall, the whole brain analyses did not reveal any corresponding right-lateral effects to the effects observed around left BA 38.

Sensor data

Of the four left hemisphere sensor groups included in the RMS analysis, the more posterior midline group exhibited activity clusters that qualitatively replicated the source results for BA 38. Most likely, these sensors were capturing the more posterior dipolar component of the flux associated with the magnetic field originating from left anterior temporal cortex (given that our sensors are axial gradiometers, the measured flux is always around the source activity, never on top of it). In these sensors (bottom of Fig. 2), when the combinatory conditions were compared to their one-word controls at the head word, a cluster of increased activity was observed for SpecGen over xptGen at 232 – 280 ms ($p = 0.05$), closely resembling the BA 38 source results. Like in the source analysis, no other comparisons between two-word and one-word conditions yielded reliable or near-reliable clusters (GenGen vs. xptGen, 259 – 278 ms, $p = 0.34$; SpecSpec vs. xptSpec, 218 – 237 ms, $p = 0.23$; GenSpec vs. xptSpec, 229 – 265 ms, $p = 0.13$).

When using the same midline posterior sensor group to test for increases for specific words over general words in non-combinatory contexts, specific words trended towards larger amplitudes both in the modifier position (221 – 263 ms after modifier onset; $p = 0.07$) and in the head position when occurring without a modifier (238–280 ms after head onset, $p = 0.08$). While in the head position, this effect may plausibly have been driven by activity within the LATL, given a small but similar temporal pole effect in the uncorrected whole brain analysis, the whole brain analysis did not yield any evidence for a LATL increase for specific over general words in the modifier position. Instead, a small increase for specific over general words was observable in superior frontal region, which may have contributed to the sensor effect.

In sum, although statistically weak, our sensor data manifested a pattern broadly compatible with the BA 38 source results, suggesting that perhaps with more subjects, even relatively crude sensor space analyses may be used to study the computations executed in the LATL.

Discussion

Composition effect as a change of concept specificity

The primary goal of this study was to investigate the interplay between the effects of single-word concept specificity and composition on the LATL, i.e., whether and how single-word concept specificity affects the composition effect. To address this, we used noun-noun compounds and manipulated the specificity of both the modifier and the
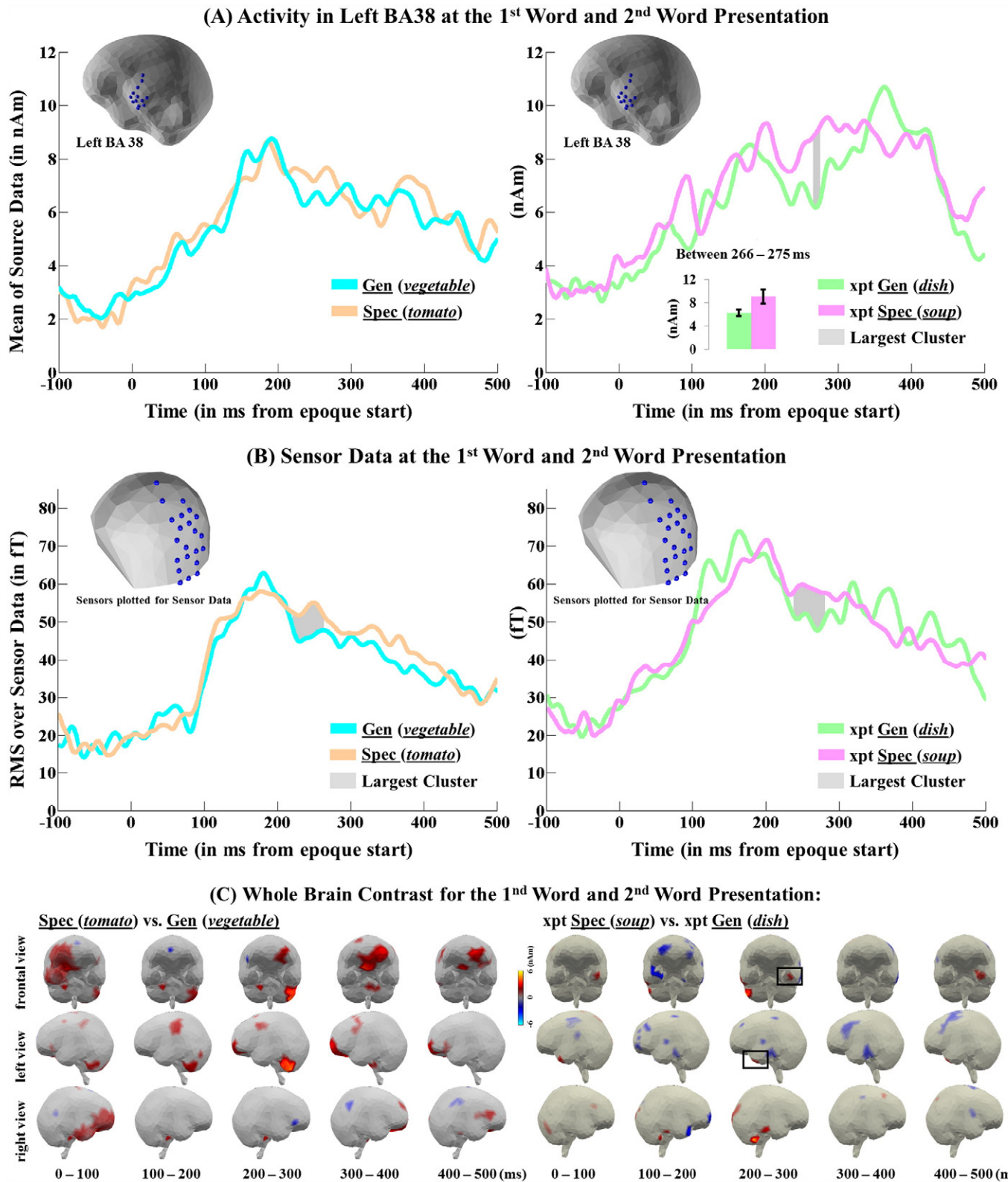
Fig. 3. The effect of single-word specificity in 1st- and 2nd-word positions. (A) Activation (in nAm) by condition in the left BA38, where 0 ms represents the onset of the 1st position word (left) and the onset of the 2nd position word (right). No clusters were found for the comparison of single word conditions at the 1st word position. A small cluster was found between 266 ms and 275 ms ($p = 0.23$) for the comparison of single word conditions at the 2nd word position. (B) Sensor data (in fT) by condition, where 0 ms represents the onset of the 1st position word (left) and the onset of the 2nd position word (right). No significant clusters were found, although in both comparisons, the largest cluster showed a pattern that a specific modifier or head word caused more magnetic field than a general modifier or head word. (C) Uncorrected whole brain comparisons at the presentation of the 1st word (left) and at the presentation of the 2nd word (right), contrasting more specific and more general modifiers (left) or head words (right).

head word. As both the ROI results and the whole brain analysis showed, increasing modifier specificity caused a larger composition effect, and in fact, the composition effect was only reliable when a more specific modifier was added to a less specific head word. A similar though statistically weaker patterns was observed in the sensor data. Crucially, the conceptual specificity of the modifier mattered for LATL activity *not* in the modifier position, when the meaning of the modifier was being accessed, but rather in the head position, when the meaning of the modifier was being integrated with the second noun. This shows that conceptual specificity mostly strongly modulates LATL activity when complex concepts are being composed, not when single concepts are being accessed.

Overall, these results are in many ways consistent with the results of several previous MEG studies on the composition effect. First, in our permutation *t*-tests, the significant composition effect was observed at 216–298 ms and in our permutation ANOVA through all the six conditions, the significant main effect of modifier specificity was obtained at 216–246 ms. These time windows conform well to the timing of LATL combinatory effects in prior comprehension studies, all reported at 200–300 ms (Bemis & Pylkkänen, 2011, 2013a,b; Westerlund & Pylkkänen, 2014). Second, the study of Westerlund & Pylkkänen (2014) showed that the LATL composition effect was larger for less specific head words. Similarly, our four permutation *t*-tests assessing the combinatory effect also showed that only by adding a specific modifier to a less specific head word could a significant composition effect be elicited.

Thus, the current findings conform to the results of previous MEG studies on the LATL composition effect and provide further evidence on what factors drive it: our results suggest that (i) the composition effect is dependent on a relatively specific modifier and (ii) when the modifier is specific, a more general head word elicits a bigger composition effect than a more specific head word. Both of these two aspects on the composition effect suggest that the composition effect is due to the difference between concepts – between a concept in isolation and the corresponding concept in a modified context. Thus, increasing the specificity of the modifier contributes to the increase of the difference. This result pattern was manifested by our significant main effect of the modifier specificity, showing that for combinatorial phrases, the LATL amplitude elicited at the head noun depends largely on the specificity of the modifier. These results fully conform to the results of Westerlund & Pylkkänen (2014) as long as their adjectival modifiers are regarded as specific modifiers, which, although harder to establish for adjectives, is a reasonable assumption.

There might be several ways to explain this prominence of modifier specificity in the processing of the whole combinatorial phrase. One possibility is in terms of diagnosticity, as suggested in Westerlund & Pylkkänen (2014). Here diagnosticity refers to how much given information contributes to distinguishing one concept from another (Smith & Osherson, 1984; Smith et al., 1988). Intuitively, by adding a modifier *red* to a noun head word *boat*, not only the modifier reduces the set of potential objects, but also it highlights the dimension of color and makes the dimension of color more prominent than other dimensions of the concept *boat* (e.g., shape, material, etc.) Thus, since the modifier could highlight a certain dimension, it would follow that the modifier plays a dominant role in the processing of the whole combinatorial phrase. Another potential explanation might be due to the word order of English: the modifier precedes the noun head word it modifies. Thus, presumably, the specificity of the modifier might be carried over in the processing of the whole combinatorial phrase. Further studies are needed to account for the prominence of modifiers shown in the processing of combinatorial phrases.

Finally, since the results of the current experiment showed that increasing the modifier specificity led to a larger composition effect, the results suggest that when elicited in the same design, the concept specificity of the modifier could be added to the composition effect in the same time window for the same brain region. Therefore, our study conforms to the view that the composition effect and the concept specificity effect may reflect the same mechanism.

Effect of single-word specificity

As regards the effect of conceptual specificity in the processing of single words in different structural positions, both our ROI and whole brain results provided subtle evidence that single-word concept specificity affected LATL amplitudes at the head words, but not at the modifiers. In other words, according to our results, only in the processing of a completed conceptual representation (here at the head words) could the single-word concept specificity effect be attested.

These results conform to prior findings on the time course and directionality of concept specificity effects. First, our concept specificity effect (i.e., the specificity effect for head words in isolation) occurred at 266 – 275 ms, consistent with the timing of previous concept specificity effects in MEG, which both for words (Westerlund & Pylkkänen, 2014) and for non-linguistic visual stimuli (Clarke et al., 2011, 2013) have occurred at about 240 – 300 ms. Second, as Fig. 3 shows, within the biggest cluster (at 266 – 275 ms), more activity was elicited in the left temporal pole for more specific concepts than for less specific concepts, as prior MEG (Clarke et al., 2011, 2013; Westerlund & Pylkkänen, 2014) and hemodynamic studies (e.g., Rogers et al., 2006) have also found.

Third, in contrast to the results of hemodynamic studies (see the literature in Section 1.1.1), the results of the current study and those of Westerlund & Pylkkänen (2014) show that the effect of single-word concept specificity is very subtle in MEG. This could be due either to the fact that we more strictly controlled for various lexical factors than previous hemodynamic studies (e.g., frequency, number of morphemes, etc.) or to the difference in imaging techniques and data analyses; for example, here we only analyzed activity time-locked to word onsets, and thus, any activity with more latency variation would have been missed. Besides, as Table 2 shows, although in general, the lexical factors of the stimuli were well matched, there were still some differences between words used in different conditions. In particular, general meaning head words had a higher mean in length and number of syllables than specific meaning head words, which might turn out to be a disadvantage in eliciting a significant single-word specificity effect.

However, for the modifier position, in the current study, both the ROI and whole brain analyses yielded no clusters at all (note that for the comparison between specific and general modifiers, the sample size is double of the sample size for the comparison between specific and general head words.) One way to understand this contrast between modifiers and head words is in terms of the semantic hub hypothesis of the LATL: according to this account, the LATL collects distributed features to form a holistic concept, and therefore, since the appropriate interpretation of modifiers has to be delayed until the head words are encountered, it would follow that the semantic hub might not treat modifiers as concepts until the head words have also been encountered. In the current experiment, participants always knew that the modifiers would be followed by a head noun, making this interpretation plausible.

Conclusion

To summarize, this study investigated how single-word specificity affects the composition effect in the left temporal pole. With a noun-noun compound design in which the specificity of nouns in both positions was manipulated, we found evidence that single-word specificity does affect the processing of a combinatorial phrase in the left temporal pole: the composition effect was reliably modulated by modifier specificity. Specifically, a reliable composition effect was only observed when a more specific modifier was combined with a less specific head noun, causing a large boost in specificity to the currently processed word. In the absence of a combinatory environment, specificity only weakly modulated LATL amplitudes (at the single word head nouns) or had no effect at all if the currently processed word was in a position where

the reader could not yet commit to any particular conceptual representation (modifier position). In sum, our findings show that concept specificity affects temporal pole activation only in positions where the conceptual representation can be completed and in particular, when this completion involves the integration of previously encountered features significantly boosting the specificity of the current item.

Acknowledgments

References

Adlam, A.L., Patterson, K., Rogers, T.T., Nestor, P.J., Salmond, C.H., Acosta-Cabronero, J., Hodges, J.R., 2006. Semantic dementia and fluent primary progressive aphasia: two sides of the same coin? Brain 129, 3066–3080.

Balota, D.A., Yap, M.J., Hutchison, K.A., Cortese, M.J., Kessler, B., Loftis, B., Neely, J.H., Nelson, D.L., Simpson, G.B., Treiman, R., 2007. The English lexicon project. Behav. Res. Methods 39, 445–459.

Baron, Sean G., Osherson, Daniel, 2011. Evidence for conceptual combination in the left anterior temporal lobe. NeuroImage 55, 1847–1852.

Baron, Sean G., Thompson-Schill, Sharon L., Weber, Matthew, Osherson, Daniel, 2010. An early stage of conceptual combination: Superimposition of constituent concepts in left anterolateral temporal lobe. Cogn. Neurosci. 1, 44–51.

Bemis, D.K., Pylkkänen, Liina, 2011. Simple Composition: A Magnetoencephalography Investigation into the Comprehension of Minimal Linguistic Phrases. J. Neurosci. 31, 2801–2814.

Bemis, D.K., Pylkkänen, Liina, 2013a. Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. Cereb. Cortex 23, 1859–1873. http://dx.doi.org/10.1093/cercor/bhs170.

Bemis, D.K., Pylkkänen, Liina, 2013b. Flexible composition: MEG evidence for the deployment of basic combinatorial linguistic mechanisms in response to task demands. PLoS One 8. http://dx.doi.org/10.1371/journal.pone.0073949.

Binder, J.R., Desai, R.H., 2011. The neurobiology of semantic memory. Trends in Cognitive Sciences 15 (11), 527–536.

Binney, R.J., Embleton, K.V., Jefferies, E., Parker, G.J., 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. Cereb. Cortex 20, 2728–2738.

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436.

Brennan, J., Pylkkänen, L., 2012. The time-course and spatial distribution of brain activity associated with sentence processing. NeuroImage 60, 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., 2012. Syntactic structure building in the anterior temporal lobe during natural story listening. Brain Lang. 120, 163–173. http://dx.doi.org/10.1016/j.bandl.2010.04.002.

Clarke, Alex, 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.

Clarke, Alex, Taylor, Kirsten I., Devereux, Barry, Randall, Billi, Tyler, Lorraine K., 2013. From Perception to Conception: How Meaningful Objects Are Processed over Time. Cereb. Cortex 23, 187–197.

Davies, M., 2008. The corpus of contemporary american english: 450 million words, 1990–2012.

Del Prato, Paul, Pylkkänen, Liina, 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.

Desgranges, B., Matuszewski, V., Piolino, P., Chételat, G., Mézenge, F., Landeau, B., de la Sayette, V., Belliard, S., Eustache, F., 2007. Anatomical and functional alterations in semantic dementia: a voxel-based MRI and PET study. Neurobiol. Aging 28, 1904–1913.

Done, D. John, Gale, Tim M., 1997. Attribute Verification in Dementia of Alzheimer Type: Evidence for the Preservation of Distributed Concept Knowledge. Cogn. Neuropsychol. 14, 547–571. http://dx.doi.org/10.1080/026432997381475.

Federmeier, K.D., 2007. Thinking ahead: The role and roots of prediction in language comprehension. Psychophysiology 44, 491–505.

Friederici, A.D., Kotz, S.A., 2003. The brain basis of syntactic processes: functional imaging and lesion studies. NeuroImage 20, S8–S17.

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.

Gainotti, G., 2006. Anatomical functional and cognitive determinants of semantic memory disorders. Neurosci. Biobehav. Rev. 30, 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, 1591–1607.

Gainotti, G., 2012. The format of conceptual representations disrupted in semantic dementia: a position paper. Cortex 48, 521–529.

Garrard, P., Carroll, E., 2006. Lost in semantic space: a multi-modal, non-verbal assessment of feature knowledge in semantic dementia. Brain 129, 1152–1163.

Garrard, P., Hodges, J.R., 2000. Semantic dementia: clinical, radiological and pathological perspectives. J. Neurol. 247, 409–422.

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, 199–212.

Grabowski, T.J., Damasio, H., Tranel, D., Cooper, G.E., Ponto, L.L.B., Watkins, G.L., Hichwa, R.D., 2003. Residual naming after damage to the left temporal pole: a PET activation study. NeuroImage 19, 846–860.

Hämäläinen, M., Hari, R., Ilmoniemi, R.J., Knuutila, J., Lounasmaa, O.V., 1993. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. Rev. Mod. Phys. 65 (2), 413.

Hodges, J.R., Patterson, K., Oxbury, S., Funnell, E., 1992. Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. Brain 115, 1783–1806.

Hodges, J.R., Graham, N., Patterson, K., 1995. Charting the progression in semantic dementia: Implications for the organization of semantic memory. Memory 3, 463–495.

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.

Humphries, C., Love, T., Swinney, D., Hickok, G., 2005. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. Hum. Brain Mapp. 26, 128–138.

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, 665–679.

Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn. Sci. 4, 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 vol. 3. JAI Press, Inc., Greenwich, CT, 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, CA, pp. 83–143.

Lambon Ralph, M.A., Patterson, K., 2008. Generalization and differentiation in semantic memory: insights from semantic dementia. Ann. N. Y. Acad. Sci. 1124, 61–76. http://dx.doi.org/10.1196/annals.1440.006.

Lambon Ralph, M.A., Pobric, G., Jefferies, E., 2009. Conceptual knowledge is underpinned by the temporal pole bilaterally: convergent evidence from rTMS. Cereb. Cortex 19, 832–838. http://dx.doi.org/10.1093/cercor/bhn131.

Lancaster, J.L., Rainey, L.H., Summerlin, J.L., Freitas, C.S., Fox, P.T., Evans, A.C., Toga, A.W., Mazziotta, J.C., 1997. Automated labeling of the human brain: a preliminary report on the development and evaluation of a forward-transform method. Hum. Brain Mapp. 5, 238–242.

Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T., 2000. Automated Talairach atlas labels for functional brain mapping. Hum. Brain Mapp. 10, 120–131.

Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (De)constructing the N400. Nat. Rev. Neurosci. 9 (12), 920–933.

Lau, E., Almeida, D., Hines, P.C., Poeppel, D., 2009. A lexical basis for N400 context effects: Evidence from MEG. Brain Lang. 111 (3), 161–172.

Maris, E., Oostenveld, R., 2007. Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190.

Mazoyer, B.M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., Mehler, J., 1993. The cortical representation of speech. J. Cogn. Neurosci. 5, 467–479. http://dx.doi.org/10.1162/jocn.1993.5.4.46.

Miceli, G., Capasso, R., Daniele, A., Esposito, T., Magarelli, M., Tomaiuolo, F., 2000. Selective deficit for people's names following left temporal damage: an impairment of domain-specific conceptual knowledge. Cogn. Neuropsychol. 17, 489–516.

Miller, Geroge A., 1995. WordNet: a lexical database for English. Commun. ACM 38, 39–41.

Miller, Geroge A., Beckwith, Richard, Fellbaum, Christiane, Gross, Derek, Miller, Katherine J., 1990. Introduction to WordNet: an on-line lexical database. Int. J. Lexicogr. 3, 235–244.

Mosher, J.C., Leahy, R.M., 1998. Recursive MUSIC: a framework for EEG and MEG source localization. IEEE Trans. Biomed. Eng. 45, 1342–1354.

Mummery, C.J., Patterson, K., Wise, R.J.S., Vandenbergh, R., Price, C.J., Hodges, J.R., 1999a. Disrupted temporal lobe connections in semantic dementia. Brain 122, 61–73. http://dx.doi.org/10.1093/brain/122.1.61.

Mummery, C.J., Shallice, T., Price, C.J., 1999b. Dual-process model in semantic priming: A functional imaging perspective. NeuroImage 9, 516–525.

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, 36–45.

Murphy, Gregory L., 2002. The Big Book of Concepts. The MIT Press.

Murphy, G.L., Brownell, H.H., 1985. Category differentiation in object recognition: typicality constraints on the basic category advantage. J. Exp. Psychol. Learn. Mem. Cogn. 11, 70–84.

Olson, Ingrid R., Plotzker, Alan, Ezzyat, Youssef, 2007. The Enigmatic temporal pole: a review of findings on social and emotional processing. Brain 130, 1718–1731.

Pallier, C., Devauchelle, A.-D., Dehaene, S., 2011. Cortical representation of the constituent structure of sentences. http://dx.doi.org/10.1073/pnas.1018711108.

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, 169–183.

Patterson, Karalyn, Nestor, Peter J., Rogers, Timothy T., 2007. Where do you know what you know? The representation of semantic knowledge in the human brain. Nat. Rev. Neurosci. 8, 976–987. http://dx.doi.org/10.1038/nrn2277.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. Spat. Vis. 10, 437–442.

Poremba, A., Malloy, M., Saunders, R.C., Carson, R.E., Herscovitch, P., Mishkin, M., 2004. Species-specific calls evoke asymmetric activity in the monkey's temporal poles. Nature 427, 448–451.

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. (Eds.), Neurobiology of Language, 1st ed Elsevier, NY (in press).

Pylkkänen, L., Marantz, A., 2003. Tracking the time course of word recognition with MEG. Trends Cogn. Sci. 7, 187–189.

Rogalsky, C., Hickok, G., 2009. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. Cereb. Cortex 19, 786–796. http://dx.doi.org/10.1093/cercor/bhn126.

Rogers, T.T., Patterson, K., 2007. Object categorization: Reversals and explanations of the basic-level advantage. J. Exp. Psychol. Gen. 136, 451–469. http://dx.doi.org/10.1037/0096-3445.136.3.451.

Rogers, T.T., Ralph, M.A.L., Hodges, J.R., Patterson, K., 2004. Natural selection: The impact of semantic impairment on lexical and object decision. Cogn. Neuropsychol. 21, 331–352.

Rogers, Timothy T., Walting, Lawrence, Hodges, John R., Patterson, Karalyn, 2005. A basic-level disadvantage for speeded category-verification. Cognitive Neuroscience Society Annual Meeting Program.

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, 201–213.

Rosch, Eleanor, Mervis, Carolyn B., Gray, Wayne D., Johnson, David M., Boyes-Braem, Penny, 1976. Basic objects in natural categories. Cogn. Psychol. 8, 382–439.

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.

Smith, E.E., Osherson, D.N., 1984. Conceptual combination with prototype concepts. Cogn. Sci. 8 (4), 337–361. http://dx.doi.org/10.1016/s0364-0213(84)80006-3.

Smith, E.E., Osherson, D.N., Rips, L.J., Keane, M., 1988. Combining prototypes: a selective modification model. Cogn. Sci. 12 (4), 485–527. http://dx.doi.org/10.1207/s15516709cog1204_1.

Snowden, J.S., Goulding, P.J., Neary, D., 1989. Semantic dementia: A form of circumscribed cerebral atrophy. Behav. Neurol. 2, 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, 2995–2999.

Tranel, Daniel, 2006. Impaired naming of unique landmarks is associated with left temporal polar damage.

Tranel, Daniel, 2009. The left temporal pole is important for retrieving words for unique concrete entities. Aphasiology 23, 867–884. http://dx.doi.org/10.1080/02687030802586498.

Van Petten, C., Luka, B.J., 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain Lang. 97 (3), 279–293.

Vandenberghe, R., Nobre, A.C., Price, C.J., 2002. The response of left temporal cortex to sentences. J. Cogn. Neurosci. 14, 550–560.

Warrington, Elizabeth K., 1975. The selective impairment of semantic memory. Q. J. Exp. Psychol. 27, 635–657.

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

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.