Adjective conjunction as a window into the LATL's contribution to conceptual combination

Eva B. Poortman¹ and Liina Pylkkänen^{2,3,4}

¹Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, 3512 JK Utrecht, The Netherlands

²Department of Linguistics, New York University, 10 Washington Place, New York, NY 10003, USA

³Department of Psychology, New York University, 6 Washington Place, New York, NY 10003, USA

⁴NYUAD Institute, New York University Abu Dhabi, PO Box 129188, Abu Dhabi, UAE

Corresponding author

Liina Pylkkänen
New York University
10 Washington Place
New York, NY 10003
USA
Tel: (212) 992 8764
liina.pylkkanen@nyu.edu

Abstract

Though a large literature implicates the left anterior temporal lobe (LATL) for combinatory operations, recent MEG studies have suggested that it is specifically involved in the composition of complex concepts, rather than syntactic or semantic composition in a more general sense. To further specify the computational contribution of the LATL, we tested whether LATL effects as observed in MEG require a situation in which features combine to form a single coherent entity representation or whether the relevant computation simply requires the attribution of features to a set but not necessarily to the same members of the set. Under the former hypothesis, the LATL would be sensitive to the number of features added to the representation of a single entity whereas under the latter account, LATL activity would reflect the total number of features integrated across different members of a set. To test this, we employed conjunctions of two adjectives whose lexical semantics were varied such that they either allowed or disallowed the attribution of their denoted properties to the same members of a set, i.e., the properties were either compatible or incompatible. The compatible properties resulted in so-called intersective and the incompatible in so-called collective readings. Our results show that the LATL tracks the number of features attributed to an individual as opposed to the number of features attributed to a set. Interestingly, the reverse pattern was found in the right ATL, demonstrating that although this region often shows parallel effects to the LATL, its functional contribution is clearly distinct.

INTRODUCTION

A fundamental aspect of human language is the productive composition of complex expressions from simpler ones. For example, a proficient speaker of English can rapidly and effortlessly combine individual words like *man* into phrases such as *tall man*, and multiple phrases into sentences such as *the tall man saw a dog*. A central goal for the neuroscience of language is uncovering the brain bases of linguistic composition, thus unpacking the specific computations underlying the composition of complex meaning. Previous research has identified the left anterior temporal lobe (LATL) as a key contributor to sentence processing in general (e.g. Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Brennan and Pykkänen, 2012), and to the basic combinatory operations of language more specifically, as for instance exemplified by the composition of small adjective-noun phrases like *red boat* (e.g. Bemis and Pykkänen, 2011, 2013a, b). More recent MEG work aimed at uncovering the particular aspects of composition that the LATL supports has suggested that it is specifically involved in the composition of complex concepts, rather than syntactic or semantic composition in a more general sense (Westerlund and Pykkänen, 2014; Zhang and Pykkänen, 2015; Del Prato and Pykkänen, 2014), in line with literature implicating the LATL for aspects of semantic memory (e.g. Baron and Osherson, 2011). The MEG work has shown that the LATL combinatory effect is sensitive to the conceptual specificity of the composing items (Westerlund and Pykkänen, 2014) and that it is absent for cases of numeral quantification (e.g., *two cups*), where the first word of the phrase does not add a feature to the noun but rather enumerates the number of tokens within a set (Del Prato and Pykkänen, 2014). It thus appears that the computation that is relevant for the LATL is some process in which a conceptual feature X is added to a representation Y.

To further specify the computational contribution of the LATL to combinatory processing, this study tested whether LATL effects as observed in MEG require a situation in which features combine to form a single coherent entity representation or whether the relevant computation simply requires the attribution of features to a set but not necessarily to the same members of the set. Under the former hypothesis, the LATL would be sensitive to the number of features added to the representation of a single entity whereas under the latter account, LATL activity would reflect the total number of features integrated across different members of a set. To test this, we employed conjunctions of two adjectives whose lexical semantics were varied such that they either allowed or disallowed the attribution of their denoted properties to the same members of a set, i.e., the properties were either compatible (*green and big*) or incompatible (*small and big*).

Prior evidence for composition and conceptual processing in the left anterior temporal lobe

Several literatures have implicated a combinatory role for the LATL. Within hemodynamic research, many studies have compared the processing of full syntactically well-formed sentences to the processing of unstructured word lists, systematically showing an increased LATL response for the sentences (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Vandenberghe et al., 2002; Humphries et al., 2005, 2006; Jobard et al., 2007; Rogalsky and Hickock, 2009; Pallier et al., 2011). Crucially, however, the processing of a full sentence is a complicated endeavor, involving a multitude of sub-processes, such as syntactic parsing of the sentence structure, semantic composition of the sentence meaning, pragmatic inferencing, and so forth. Thus several different hypotheses regarding the function of the LATL remain consistent with the sentence vs. list findings.

Later work using MEG has aimed to narrow down this hypothesis space using very minimal experimental designs (Bemis and Pylkkänen, 2011, 2013a, b). These studies moved away from the sentence vs. list paradigm, and instead compared simple adjective-noun combinations (e.g. *red boat*) to one-word stimuli (e.g. *boat*) or two-word lists (e.g. *cup, boat*). Results of these studies consistently showed more activity in the LATL for the adjective-noun condition than for their non-combinatory controls. Thus these findings suggest that the LATL effects observed in sentence processing studies reflect basic combinatorial operations as opposed to other components of sentence processing. A remaining question within this line of work, however, is whether the LATL is responsible for syntactic or semantic composition, since the processing of adjective-noun combinations clearly involves both.

In answer to this, subsequent research has shown that the LATL is unlikely to reflect any general form of either syntactic or semantic composition, but rather, it appears more specifically involved in the composition of some type of complex concepts (Westerlund and Pylkkänen, 2014; Del Prato and Pylkkänen, 2014), as also suggested by fMRI results from visual categorization tasks on complex concepts and their constituents concepts (Baron & Osherson, 2011; Baron et al., 2010). This hypothesis is in line with literature on the ‘hub and spoke’ model of conceptual knowledge (e.g. Patterson et al., 2007; Lambon Ralph et al., 2010), which characterizes the LATL as a semantic, amodal hub connecting a widely distributed neural network of areas that each contain modality-specific features of concepts (including sensory, motor and verbal areas). The hub serves as an additional layer that integrates the different types of conceptual information in order to build an amodal representation of a concept. Consistent with this, ATL atrophy correlates with a disorder in conceptual knowledge called Semantic Dementia (SD), in which categorization performance is modulated by concept specificity: specific concepts, with more distinctive features, are more affected than less specific ones (Warrington, 1975; Hodges et al., 1995; Done & Gale, 1997; Rogers et al., 2004, 2006). This pattern can be predicted by the

hub and spoke model if it is assumed that the hub is responsible for categorization based on featural overlap. To investigate the relationship between combinatorial and conceptual specificity effects in the LATL, Westerlund and Pykkänen (2014) compared low and high specificity nouns (e.g. *boat* vs. *canoe*) in both combinatorial and non-combinatorial contexts (e.g. *blue boat/blue canoe* vs. *xhsl boat/xhsl canoe*), finding that low specificity nouns exhibited a larger amplitude increase as a result of adjectival modification than high specificity nouns, i.e., the LATL showed a larger combinatorial effect for lower specificity items. Thus it is possible that the specificity boost provided by adjectival modification was larger when the adjective modified a less specific noun, a generalization also supported by a subsequent finding where more specific modifiers (such as *vegetable* in *vegetable soup*) were found to elicit larger combinatorial effects on the noun than less specific ones (Zhang & Pykkänen, 2015). These findings are unpredicted by functional hypotheses associating the LATL to syntactic or semantic composition in general terms, as these computations should not be sensitive to conceptual specificity, and instead conform to an account in which the LATL more specifically reflects the construction of complex concepts. Of extant more general models of the neurobiology of language, these findings fit quite naturally with Bornkessel-Schlesewsky et al.'s (2015) framework where the role of the ATL is conceived of as a type of feature combiner within a ventral stream of auditory object recognition, assumed to have evolved to serve language processing more generally (see also Bornkessel-Schlesewsky & Schlewsky (2013)). A core assumption of their model is that anterior temporal cortex combines features in an order-insensitive way (Bornkessel-Schlesewsky et al., 2015), a hypothesis partially supported by MEG data from the minimal composition paradigm. Specifically, Bemis & Pykkänen (2013b) tested whether the early combinatorial effect in the LATL that is also the target of the current investigation would be present for stimuli in which adjectives and nouns are presented in the reverse order (*boat red*), ungrammatical in English in the absence of further context. Results showed that this was the case but only if the experimental task required subjects to combine the meanings of the words. Under task demands that did not require composition, only the grammatical sequences elicited a LATL combinatorial effect. Thus it appears that the LATL combinatorial function at ~200-250ms is computationally able to combine reversed order stimuli, but this is subject to top-down control. In contrast to this work, which addressed the automaticity of the LATL response, the focus of the present study was to add precision to our understanding of what variety of conceptual combination is in fact relevant for the LATL.

Current study: Intersective vs. collective conjunction as a window to LATL function

In the studies summarized above, composition or conceptual combination was always achieved by modifying nouns with adjectives, as in *red boat*. Another way to add features to a concept is via

conjunction: *this box is red and large*. The extant literature on the LATL would strongly predict the LATL to be sensitive to this type of conceptual combination. However, conjunctions do not necessarily add features to a single individual, instead, they may also distribute the introduced features to different members of a set. Consider a room full of tall blonde girls, some of whom are Dutch and some of whom are Finnish. This situation could be truthfully described with the sentence *these girls are Dutch and Finnish* but crucially, the two predicates of the conjunction, *Dutch and Finnish*, would not add features to the same individuals within a set. In contrast, the sentence *these girls are tall and blonde* is generally interpreted as attributing tallness and bloneness to all members of the set. We shall call the former interpretation a **collective** interpretation, as it attributes both Dutchness and Finnishness to the set of girls collectively but not to the individuals (e.g. Krifka, 1990; Heycock and Zamparelli, 2005; Winter, 2001). The latter interpretation will be called **intersective** reflecting the so-called Boolean analysis of conjunction, according to which it behaves as set-theoretic intersection (Keenan and Faltz, 1985; Partee and Rooth, 1983), i.e., a sentence such as *the girls are tall and blonde* is considered to be true if and only if every girl that is being referred to is in the intersection of the set of tall entities and the set of blonde entities, i.e. if each girl is both tall and blonde.

If the notion of “conceptual combination” that is relevant for the LATL involves the combination of features to form a single coherent entity representation (which of course could be a plurality as in *these girls are tall and blonde*), then only intersective interpretations should engage the LATL. This is what a hub and spoke model most naturally predicts: Since intersective conjunction adds features to the same individuals, it serves to further specify those individuals’ representation, i.e., *these girls are tall and blonde* results in a more specific girl-representation than the single property predication of *these girls are blonde*. In contrast, the collective interpretation of *these girls are Dutch and Finnish* does not result in a more specific girl-representation than the single property predication of *these girls are Finnish*, as in both cases any girl within the set gets attributed one nationality-feature (putting aside dual citizenship interpretations, which would be intersective). Thus under this hypothesis, only intersective conjunction should elicit increased LATL activity as compared to one-property predication. However, if the computation relevant for the LATL simply reflects the attribution of features to a set but not necessarily to the same members of the set, then both intersective and collective interpretations should engage the LATL. This latter hypothesis would attribute a more general combinatory role for the LATL and indeed a finding where the LATL shows sensitivity to both intersective and collective conjunction would be compatible with many functional hypotheses, including syntactic and semantic composition in a very general sense.

The goal of this work was to distinguish between these two hypotheses with a manipulation where participants judged the truthfulness of sentences involving either collective or intersective

conjunctions. Though conjunctions are of course surface ambiguous between intersective and collective interpretations, it has been shown that the less compatible the two properties described by the members of the conjunction are, the more likely a collective interpretation is (Poortman, 2014). In this work subjects rated the acceptability of plural sentences with two conjoined verbs in a split situation in which one verb applied to only two individuals and the second verb applied to two other individuals (thus conforming to a collective interpretation of the sentence). The verb pairs that were included in the experiment systematically varied in the degree to which they were compatible, i.e. could apply to one individual simultaneously, as independently measured in a rating task on a scale. Findings revealed that the degree to which the verb pairs were incompatible correlated with the degree to which the sentences were accepted, thus the degree to which collective interpretations were available. For example, *sitting and standing* was independently rated to be a highly incompatible pair, and consequently the sentence *the boys are sitting and standing* showed high acceptability in a split situation in which two boys were sitting and two boys were standing. On the other hand, *sitting and reading* was judged perfectly compatible, and when used in a plural sentence (*the boys are sitting and reading*), it showed low acceptability given the split situation.

Our design built on these findings and contrasted sentences that are syntactically identical, but differed with respect to the conceptual combination of features that is required. Specifically, we contrasted sentences of the form *the X are P1 and P2*, where *P1* and *P2* either allowed intersection (a color and a size adjective) or strictly disallowed it (2 size adjectives that were antonyms). The predicates used in the experiment were all adjectives since, unlike verbs, adjectives unambiguously allow the construction of strictly incompatible pairs (e.g. *big and small*) and hence exclude the possibility of conceptual combination of those two predicates. Participants first viewed displays of objects and then indicated whether the subsequent sentence was a truthful description of the display. MEG analysis targeted the last adjective of the sentences, which was either the second adjective of the conjunction conditions or the sole adjective of our single-adjective baseline condition (Fig. 1). It is important to note that our one-adjective baseline condition was itself a sentence, i.e., not noncombinatory like e.g., the one-word controls in several prior MEG studies on this topic (e.g., Bemis & Pylkkänen, 2011, 2013a, b). Thus our control condition was itself predicted to engage combinatorial processing in the LATL to some extent, though, by hypothesis, less than the presentation of two adjectives. This consequently sets up the expectation that the one vs. two adjective conditions could diverge from each other more subtly in the LATL than in the prior studies where the baseline was entirely noncombinatory.

In addition to the LATL, we included as a region of interest (ROI) its right hemisphere homologue, as anterior temporal lobe (ATL) effects have often been bilateral (Mazoyer et al., 1993; Stowe et al., 1998; Friederici et al., 2000; Humphries et al., 2001; Bemis & Pylkkänen, 2011; Leffell et al., 2014), with significant debate addressing possible functional divisions between the two hemispheres (e.g.,

Humphries et al., 2005; Rice, Ralph, & Hoffman, 2015). Thus as an additional goal to our study we aimed to test the extent to which the pattern of ATL effects observed in the left hemisphere was mirrored in the right.

METHODS

Participants

A total of fifteen right-handed native speakers of English participated in the study (13 female; average age 23.5 years). All participants were non-colorblind and had normal or corrected-to-normal vision. Before participation, they all signed an informed consent form.

Experimental design and stimuli

Our stimuli consisted of visual displays followed by simple sentences which were judged for veridicality as descriptors of the display. Though somewhat artificial, our aim was to mimic naturalistic conditions where linguistic expressions are interpreted in a given context, here provided by the display. Test sentences contained adjective conjunctions, with the adjectives either allowing an intersective interpretation or forcing a collective one. The conjunction contexts were also compared to a baseline condition in which sentences contained only one adjective. The three conditions and the structure of the trials are shown in Fig. 1. The target of MEG analysis was always the last (or sole) adjective of the sentence.

The intersective conjunction condition consisted of sentences containing a plural subject and two conjoined intersective adjectives: a color and an adjective describing size (e.g. *The hearts are green and big*) whereas the collective conjunction condition consisted of sentences with two conjoined adjectives that force a collective reading: two incompatible size adjectives (e.g. *The hearts are small and big*). The condition without conjunction consisted of sentences with only one adjective. Each sentence contained a period at the end, which appeared on the screen together with the final (or sole) adjective (see Fig. 1). To make sure participants were paying attention, they were asked to indicate for each sentence whether it was true or false for a colored image of four shapes that preceded it. To count as matching in the intersective condition, both adjectives (color and size) had to apply to each individual shape in the image. To count as matching in the collective condition, one adjective had to apply to two of the individual shapes while the other applied to the remaining two shapes. In the trials without conjunction, only the shape of the objects was relevant for the decision.

Each trial contained a context image, a test sentence containing either one or two adjectives (depending on condition), and an image depicting two fingers plus “yes” and “no” buttons (Fig. 1.) prompting the participant to perform a task (to avoid quick button presses in response to our critical and sentence-final word). A total of six size adjectives (*big, small, thin, thick, long, short*) and six color adjectives were used (*red, blue, green, white, black, pink*) in the conjunctions. The critical sentence-final adjective was kept constant across conditions as shown in Fig. 1. Collective conjunction trials were formed by replacing the first conjunct of the intersective conjunction trial (e.g. *green and big*) with the antonym of the second conjunct (yielding *small and big*). The collective readings could only be forced with combinations of two incompatible size adjectives and thus all collective trials involved two size adjectives (a conjunction of two color adjectives can always in principle be interpreted intersectively, i.e., *pink and green*, for example, can be taken to mean ‘partly pink and partly green’). Trials without conjunction were created by removing the first conjunct (yielding *big*); since the last word of the sentence occurred with a period, participants knew to not expect a second conjunct on these trials. Predictability of the second conjunct in the collective conjunction condition was controlled for by adding a filler item to each triplet that contained the two intersective adjectives in reversed order (e.g. *big and green*).

For the noun position, twenty one-syllable nouns were used (*bag, boat, bow, cane, car, cross, cup, glass, hand, heart, house, key, lamp, leaf, lock, note, plane, shoe, star, tree*). Target images each consisted of four shapes, each unambiguously depicting one of the noun concepts, but manipulated in color and/or physical dimension.

Participants each viewed 384 trials in 12 stimulus blocks: 96 of each of the three trial types plus 96 filler items with adjectives in reversed order. In half of the trials for each condition, the target image matched the sentence. In the conjunction conditions, half of the mismatch trials could be judged as a mismatch already on the first conjunct while the other half could only be judged as a mismatch on the second conjunct. Predictability of the final (or sole) adjective in match trials was equal across all conditions (namely 100%), assuming participants generally predict for a sentence that matches the context picture.

In each condition, each of the 20 nouns was used six times, three times in matching trials and three times in mismatching trials. The order of trials within stimulus blocks was randomized and constructed separately for each participant. Only the match trials were included in the analysis.

Procedure

Prior to taking part in the experiment, participants received written instructions on the task and saw examples of the types of images that they would see in the experiment. They then performed a practice

session outside of the MEG room. This session, unlike the actual experiment, provided feedback in order to ensure participants' comprehension of the task. All practice trials were distinct from experimental trials.

Before starting the MEG recording, participants' head shapes were digitized using a Polhemus Fastscan three-dimensional laser (Polhemus Inc, Vermont, USA) in order to determine the location of five marker coils that were placed across participants' heads, whose positions were measured before and after the recording. During the analysis, the digitized head shapes were used to constrain the source localization around a participant's head with respect to the MEG sensors.

During the experiment, participants lay down in a dimly lit magnetically-shielded room. They performed the task in 12 separate blocks, with the option of taking breaks between blocks. The order of blocks was randomized for each participant.

MEG data were collected using a whole-head 157-channel axial gradiometer system (Kanazawa Institute of Technology, Tokyo, Japan) with a 1000 Hz sampling rate and a low-pass filter at 200 Hz and a notch filter at 60 Hz. The screen on which the stimuli were presented was situated approximately 50 cm from the participant's eye. During each trial, the target image was presented for 600 ms. After that, words were presented one by one for 300 ms each, and were each followed by a 300 ms blank screen. An image prompting the participant to make a decision appeared at the end of each trial and remained onscreen until the participant made a decision (Figure 2).

Pre-processing of MEG data

The raw continuously recorded MEG data were noise-reduced using the Continuously Adjusted Least Squares Method (CALM; Adachi et al., 2001). Due to urban noise, the data were then high-pass filtered using a Hamming window with a cut-off frequency of 1 Hz and a filter width of 2 s. Artifacts were manually removed from trials (from 200 ms before onset of the final adjective to 1000 ms after onset) with visual inspection by rejecting those with amplitudes exceeding ± 3000 fT or containing eyeblinks.

The data were then averaged for each condition, from 200 ms pre stimulus to 1000 ms post stimulus (i.e., the sentence final adjective), with averages low-pass filtered at 40 Hz and baseline corrected from -100 to 0 ms prior to source analysis. Separate distributed L2 minimum norm source estimates were created for each condition average (distinguishing between match and mismatch items) for each subject, using BESA (version 6; MEGIS Software, GmbH, Gräfelfing, Germany). Next, we performed region-of-interest (ROI) analyses on the averaged data (match items only), focusing on the left and right anterior temporal lobes during the processing of the final adjective in the sentence. Secondly, we also performed a whole-brain source analysis to ensure that activity seen in the ROI analyses indeed

reflected activity in the LATL or RATL and exclude the possibility of spill-over effects from adjacent areas.

In addition to analysing the activity elicited by the sentence-final adjective in all conditions, for the conjunctive conditions we also analysed the activity elicited by the first conjunct and the word *and*, in order to provide a more comprehensive processing profile of the entire conjunction and to determine whether differences between the two types of target adjectives might already be present earlier in the conjunction. Since very long epochs result in more lost trials due to artifacts, we created separate averages capturing the ‘first adjective + *and*’ region, starting at 200ms before the onset of the first adjective and lasting till 1000ms after (i.e., till 400ms after the onset of *and* or 200ms before the onset of the second adjective). The analysis parameters for these pre-target averages were identical to those of the sentence-final adjectives.

Data analysis

Behavioral responses

Accuracy and reaction times (RTs) of participants’ behavioural responses were analyzed with repeated measures ANOVAs with Condition as the within-subjects factor (3 levels: intersective, collective, no-conjunction).

ROI analysis

The ROI analyses focused on the left anterior temporal lobe (LATL), specifically Brodmann area (BA) 38, which in our prior studies has been the most consistent locus of conceptually modulated LATL composition effects (Zhang & Pylkkänen, 2015; Del Prato & Pylkkänen, 2014; Blanco-Elorrieta & Pylkkänen, 2016). The right anterior temporal lobe (RATL, BA 38) was treated as an additional ROI, as motivated above. The Talairach daemon was used to automatically assign Brodmann area labels to the 713 sources on the smooth BESA cortex (Lancaster et al., 1997, 2000) and the sources within left BA 38 were averaged together to yield a single activity time course for this region and similarly for the right BA 38. The analyses were time-locked to the presentation of the final adjective in the sentence.

Differences in ROI activity were assessed with cluster-based permutation t-tests (Maris & Oostenveld, 2007) comparing the conjunction conditions to each other as well as to the non-conjunctive baseline. Analyses were conducted at 100 - 300ms after presentation of the target adjective, as previous studies have indicated that composition-related activity occurs at 200 - 250ms (Bemis & Pylkkänen 2011, 2013a, b). The thresholds for initial cluster selection were also adopted from prior MEG studies on the LATL (Bemis & Pylkkänen 2011, 2013a, b; Westerlund & Pylkkänen 2014; Del Prato & Pylkkänen

(2014): specifically, all waveform separations lasting at least 10 time samples at $p = .3$ were first grouped into clusters and of these, the one with the largest summed t -value within the analysis interval (100 - 300ms) was entered into 10,000 permutations, yielding a corrected p -value ($\alpha = .05$) that represented the proportion of times that the tests on the random partitions yielded a test statistic higher than the actual observed statistic. Finally, the p -values across our two ROIs were FDR corrected (Genovese et al. 2002) with a criterion value of 0.05.

In the pre-target region, this same analysis was performed at 100-300ms after the onset of the first adjective of the conjunctive conditions, aimed at assessing the combinatory effect reported in prior literature. Since the general goal of the pre-target analysis was to reveal any baseline issues that might affect the interpretation of the results on the target adjective, we intentionally “double-dipped” and ran additional post-hoc permutation tests on any waveform difference that had the appearance of a possible effect. This resulted in two additional analyses at 400 - 450ms and at 700 - 1000ms after the onset of the first adjective. As reported below, none of these analyses revealed reliable differences.

Whole-brain analysis

In order to ensure that the ROI analyses in fact reflected activity localized to the LATL or RATL (as opposed to spill-over activity from neighboring regions), liberally thresholded uncorrected whole-brain contrasts were generated for each pair-wise comparison in order to visualize the spatial extent of the effects revealed in the ROI analysis. We performed paired t -tests in each source on the smooth BESA cortex for a time window of 0 - 700 ms from the onset of the sentence-final adjective, and plotted activity on the standard BESA brain if a source showed an effect at the $p < 0.05$ level and was surrounded by at least 2 spatial and 2 temporal neighbors that were also reliable at the 0.05 level.

RESULTS

Behavioral data

Accuracy on average was high, with at least 90% correct responses per condition (Collect M = 94% [SD = 6%]; Intersect M = 92% [7%]; NoConj M = 90% [5%]). Repeated measures ANOVAs were performed on both accuracy and RT data with Condition as the within-subjects factor (3 levels: interjective, collective, no-conjunction). The ANOVA on accuracy showed a main effect of Condition ($F(2, 28) = 8.99$, $p < .01$), with planned pairwise comparisons revealing that only the collective and the no-conjunction conditions significantly differed from each other ($p < .001$), with higher accuracy for the collective condition. Similarly, we found a main effect of Condition for reaction time ($F(2, 28) = 24.02$, p

< .001). Here, pairwise comparisons revealed two significant differences: reaction times were longer in the no-conjunction condition compared to the collective condition ($p < .01$; NoConj M = 784ms [SD = 28ms]; Collect M = 633ms [26ms]) and compared to the intersective condition ($p < .001$; Intersect M = 600ms [25ms]). These results are surprising, since the task in the no-conjunction condition required participants to check only one feature compared to two features in the other two conditions. This may be an effect reflecting surprisal, since throughout the whole experiment the no-conjunction cases comprised only a quarter of the stimulus material. Thus crucially, any increase in neural activity for the collective or intersective conditions compared to the no-conjunction condition should not reflect increased effort, since subjects were on average more accurate and faster in responding to the two-adjective conditions. Moreover, the collective and intersective conjunction conditions did not differ in accuracy or reaction time from each other. Thus again, any effects between these two conditions are unlikely to be due to differences in general effort required in the task.

ROI data

Sentence final adjective

The cluster based permutation t-tests revealed significant clusters of time points both within the LATL and the RATL, although the result patterns in the two regions were very different (Figure 2).

In the LATL, we found a significant cluster of greater activity in the intersective condition compared to the collective condition from 162 to 275 ms ($p = 0.0119$; average $\pm$ SEM: collect, 12.70 ± 1.57 nAm; intersect, 16.49 ± 1.91 nAm). When comparing the intersective to the no-conjunction condition, a cluster of increased activity in the intersective condition was found in the LATL from 181 to 232 ms, though it did not survive correction for multiple comparisons ($p = .12$; intersect, 16.30 ± 2.59 nAm; no-conj, 12.73 ± 1.90 nAm). The comparison between collective and no-conjunction conditions showed no evidence for an increase for conjunction. Instead the data trended towards the opposite pattern, with a non-significant cluster ($p = .61$) of increased activity for the no-conjunction condition (13.33 ± 2.49 nAm) over the collective condition (9.57 ± 1.52 nAm) occurring at 115 - 142 ms. Thus our LATL results strongly indicated a contrast in the LATL's involvement in intersective vs. collective conjunctions, with intersective but not collective interpretations increasing LATL amplitudes.

Within the RATL, the comparison between the intersective and collective conditions revealed the opposite pattern from the one observed in the LATL, with increased amplitudes for collectives over intersectives. The duration of the observed effect lasted until the end of our analysis interval and therefore we lengthened the interval to 100 - 450ms in order to capture the entire cluster, which onset at 274 and lasted until 422 ms (collective: 21.79 ± 3.20 nAm; intersective: 15.01 ± 1.94 nAm). However, although

sustained, the cluster was only near-significant after multiple comparisons correction ($p = .0789$). Consistent with this profile, a highly significant increase was observed for collectives when compared to the no-conjunction condition ($p = 0.0184$, collect 17.28 ± 2.18 nAm; no-conj, 10.90 ± 1.25 nAm). This effect also lasted through the entire analysis window (100 – 300ms) and when the extent of the cluster was explored by lengthening the window, the cluster persisted from 0 to 516 ms ($p = 0.0092$, collect 17.32 ± 2.05 nAm; no-conj, 10.69 ± 1.05 nAm; analysis window: 0 – 600ms). Finally, similarly to the pattern in the LATL, the intersective condition elicited larger amplitudes than the no-conjunction condition though this effect remained marginal ($p = 0.09$ for 164–229ms, intersect, 15.54 ± 2.14 nAm; no-conj, 10.98 ± 1.36 nAm). In sum, RATL amplitudes were robustly increased for the collective interpretations and weakly so for the intersective ones.

First conjunct and “and”

For the conjunctive conditions, the time course of LATL and RATL activity was also assessed in the pre-target region, covering the first adjective and the conjunctive *and*, as shown in Fig. 3. We searched for reliable clusters in a pre-defined time-window of 100-300ms, aimed at targeting the combinatory effect reported in prior literature, as well as in two time-windows defined on the basis of the visual appearance of the waveforms; recall that our aim was to reveal any possible baseline issues that might affect the interpretation of our main result. The latter two analyses focussed on possible LATL increases for collectives over intersectives at ~400ms and for intersectives over collectives at ~800ms (top left panel of Fig. 3). None of these analyses yielded reliable differences between the conditions, despite being specifically tailored to reveal such effects (the lowest uncorrected p-value from the tests was 0.3). Thus this analysis suggests that the differences at the sentence-final adjective in fact reflect computations occurring specifically at this adjective.

Whole brain analysis

Paired t-tests over the whole brain were conducted for each pairwise comparison (Fig 2 bottom). The contrast between intersectives and collectives showed a clear activity increase for intersectives in the left temporal pole between 100 and 300ms, consistent with the ROI analysis. Other than this, there were no obvious increases for intersectives over collectives until much later in the time course (around 600ms). The collectives, however, elicited a robust activity increase over intersectives in the right hemisphere at 300 - 500ms, centered on the right temporal pole as captured by the ROI analysis but also spreading posteriorly along the ventral cortical surface.

The comparison between intersectives and no-conjunction controls showed weak and somewhat diffuse increases for intersectives along both temporal poles, again consistent with the ROI analysis in which both temporal poles trended towards larger amplitudes for intersectives in this comparison.

Finally, collectives showed a robust activity increase in the right temporal pole as compared to the no-conjunction controls, as also observed in the ROI analysis. Like in the ROI analysis, left anterior temporal cortex showed no evidence for an increase for collectives in the uncorrected whole brain contrast either, further corroborating the conclusion that the LATL's computational role does not include the interpretation of collective conjunction.

DISCUSSION

This study aimed to move beyond the general notion of conceptual combination to more precisely characterize the computational contribution of the left anterior temporal lobe in semantic processing. Given that prior studies have already demonstrated that combinatory activity localizing in the LATL at 200-250ms is sensitive to the conceptual specificity of the composing items (Westerlund & Pykkänen, 2014; Zhang & Pykkänen, 2015), we assumed that this activity does not operate at the syntactic level but rather constructs some type of conceptual representations. More specifically, the hypothesis that the LATL reflects conceptual combination assumes that its amplitude reflects the addition of features into conceptual representations. Here we aimed to sharpen this by testing whether LATL activity reflects the number of features added to an individual or the number of features added to a set. By creating a situation where the members of a set were attributed fewer features than the set collectively, we were able to tease these two hypotheses apart: in our baseline condition, one feature was attributed to each member of the set; in our intersective condition, two features were attributed to each member, whereas in our collective condition, two features were attributed to the set but each individual only received one. Our results clearly indicated that LATL amplitudes are only enhanced when the added features apply to the same individuals: distributing the features across a set did not increase this activity.

We also observed that the LATL increase elicited by intersective conjunctions is at least numerically more robust in comparison to the collective conjunction as opposed to when compared to the no-conjunction control condition, the latter contrast only eliciting a marginal effect. This was not predicted but could be due to the fact that at the single adjective in the no-conjunction condition, LATL relevant composition is likely to occur in a rather parallel way as on the second conjunct of the intersective condition. Specifically, in both positions, the property named by the currently processed adjective is attributed to all the objects described by the subject of the sentence. This parallelism could have made the Intersective vs. No-conjunction contrast a relatively weak one and in fact in most of our

prior studies, LATL combinatory effects have been elicited against baselines that are entirely non-combinatory (e.g., Bemis & Pykkänen, 2011, 2013ab, Brennan & Pykkänen, 2012; Westerlund et al., 2015). This however remains speculative as with increased power, the difference between the two effects (intersective vs. no-conjunction and intersective vs. collective) could of course have flattened.

Most importantly, our findings substantially narrow down the possible hypothesis space on the computational contribution of the LATL: no account in terms of general syntactic or semantic composition can explain these results nor any general conceptualization of conceptual combination in terms of “the construction of complex concepts from simple constituents” (Baron & Osherson, 2011). Instead, the LATL does something more specific than this, potentially limited to the construction of more specific representations of individuals within a discourse model.

If this hypothesis is on the right track, future research should then target the question: what counts as a relevant type of “individual” for the purpose of LATL computation? Theories of natural language ontology can provide some guidance for how to ask this question; for example, in some theories of formal semantics, events are treated as a type of individual within the compositional system (Davidson, 1967; Parsons, 1990) whereas in others they are not (e.g., Montague, 1970; Heim & Kratzer, 1998). In the former type, so-called event-based theories, verbs are treated as properties of events and most nouns as properties of individuals, resulting in a system in which events and individuals function as similar type of variables in the combinatory system. For example, the hypothesized interpretation of *red square ball* could be paraphrased as “an individual such that it is a ball, red and square” and the interpretation of a verb phrase such as *stab Brutus in the back* would similarly proceed as “an event such that it is a stabbing, has Brutus as an affected participant, and is in the back” (Parsons, 1990). Thus in this type of theory, the direct object *Brutus* ends up as a concatenated property in the event description, similarly to modifiers such in the back or *red* and *square* in the noun phrase. A relevant empirical question then is: does the construction of event descriptions engage the LATL similarly to the construction of entity descriptions, such as *red ball*? Westerlund et al. (2015) addressed this in a design targeting a possible computational distinction between modification and argument saturation, with results showing that the composition of phrases such as *eats meat* engages the LATL in a similar fashion to adjective noun combinations. Thus, if the LATL constructs representations of individuals, events also count as such individuals.

In broad terms, our results allow an interpretation along the lines of a ‘hub and spoke’ model of conceptual knowledge, as discussed above (e.g. Patterson et al., 2007; Lambon Ralph et al., 2010). Many studies have identified the anterior temporal lobes as candidates for the localization of this hub, by showing that damage to these regions causes Semantic Dementia, which in turn is shown to affect categorization performance depending on the specificity with which items must be recognized: specific concepts are more affected than less specific ones (Warrington, 1975; Hodges et al., 1995; Done & Gale,

1997). In the current study, the intersective conditions required the combination of two distinguishing features such as *big* and *green*, resulting in a more specific concept (i.e. *big green hearts*) than in the collective conditions, which do not require such integration of features. Instead, in the case of for example *big and small hearts*, two higher-level concepts are distributed over entities. These concepts are more robust in the sense that they are distinguishable by several different features, resulting in a less distinctive classification. Consequently, the intersective conditions, with more specific concepts like *big green hearts* are expected to engage the semantic hub more than the collective conditions, because there are more overlapping representations to differentiate. Simply put, the hub needs to instantiate the ‘heart’ representations more exactly for *big and green hearts* (in terms of both shape and color), than for *big and small hearts* (where it’s merely necessary to find ‘heart’ representations that are sufficiently like big hearts or small hearts in order to activate the concept). Consequently, the increased LATL activity in the intersective over the collective conditions might reflect the difference in specificity of concepts that are being categorized in these conditions – following the predictions of the hub and spoke model. This result is in line with neuroimaging evidence that shows more LATL activity for more specific image categorization (for example in case of subordinate categories like *robin*) than for basic level or domain general level categorization (for example in case of *bird* or *animal*) (Rogers et al., 2006). It is also in line with Westerlund & Pyllkanen’s (2014) MEG results on small adjective phrases, which show that LATL activity increases proportional to the change in concept specificity, as well as the findings that their paper was based on, namely that adjective-noun constructions like *blue boat* lead to more LATL activity than simple nouns like *boat* (Bemis and Pyllkanen, 2011, 2013a, b).

The hypothesis that the LATL constructs representations of individuals also has some similarity to another extant hypothesis about LATL function, the so-called “unique entities” account. This hypothesis arises from findings that the LATL is preferentially engaged by famous or personally familiar stimuli (Gorno-Tempini & Price, 2001; Grabowski et al., 2001) and that ATL damage can lead to specific deficits for famous/familiar entities (for a review, see Gainotti, 2007). Uniqueness can of course be thought of as a form of specificity and thus the unique entities account is not necessarily incompatible with wider hypotheses implicating the LATL for semantic processing. However, so far no studies have aimed to tease uniqueness and specificity apart (Wong & Gallate, 2012). A further empirical finding converging on the general notion that the LATL builds representations of individuals comes a decoding study showing that pattern-classification algorithms can be used to decode the detection a target’s identity in the LATL, with identity information in LATL being specifically predicted by the temporal convergence of shape and color codes in early visual regions (Coutanche & Thompson-Schill, 2014).

A valuable aspect of the current work that is absent in related hemodynamic and neuropsychological research is the ability to focus on specifically timed LATL activity, allowing for an

investigation of the specific processing stage that occurs at ~200-250ms. As regards this activity, the generalization that is emerging is that it systematically participates in the composition of complex phrases in a semantically sensitive (Westerlund & Pylkkänen, 2014; Zhang & Pylkkänen, 2015), cross-modal (Bemis & Pylkkänen, 2013a; Pylkkänen et al., 2014) and task-independent fashion (Bemis & Pylkkänen, 2013b). The current result shows that within syntactically parallel expressions, it is the semantic richness of the individuals within a set as opposed to the set collectively that modulates activity in the LATL.

Interestingly, the RATL, which was not primarily part of our hypotheses but is often thought to be responsible for similar processes as the LATL, in fact showed a very strong opposite effect from the LATL. In the RATL, the collective conditions led to significantly more activity than the one-adjective baselines and trended towards larger amplitudes than the intersectives. This is of theoretical interest since the left and right ATLs often show similar effects (Ferstl et al., 2008; Lambon Ralph et al., 2009, 2010; Visser et al., 2010), including in our own MEG work (Bemis & Pylkkänen, 2011; Leffell et al. 2014), and thus the respective functional contributions of the ATLs have been non-trivial to characterize. However, in light of more general proposals about the distinct roles of the left and right hemispheres in language processing, our hemispheric difference is less surprising. The left hemisphere is typically thought of as the faster, more local processor whereas the right hemisphere contributes to more global computations (e.g., Beeman & Chiarello, 1998). Most relevantly, there is evidence that the right hemisphere contributes to the establishment of global coherence while the left hemisphere computes local coherence relations (St. George et al., 1999). Our results could be predicted from this hypothesis: the collective readings are in a sense locally incoherent, containing a conjunction of contradictory adjectives which only become coherent when distributed to distinct members of the set named by the subject of the sentence. In contrast, the intersective conjunctions are locally coherent as they describe two mutually compatible properties. Additional evidence for a right hemisphere role in processing antonymic relations comes from neuropsychological data on right hemisphere patients who are especially bad on tests measuring their sense of opposition (Caramazza et al., 1976; Gardner et al. 1978).

The timing profiles of the two ATL effects also fit the general idea that the left hemisphere performs faster and more local computations than the right hemisphere: in the left ATL, intersectives elicited larger amplitudes than collectives at 200-250ms while the reverse, though statistically weaker, pattern obtained in the right ATL roughly one hundred milliseconds later (~300-400ms). In comparison to the one-adjective baseline though, the right ATL increase began very early, in fact at the onset of the stimulus (and thus presumably even earlier). Most likely this reflects anticipatory activity reflecting a prediction for a collective reading: whenever the first adjective truthfully described only a subset of the just viewed objects, the only conjunction that could match the picture was a collective one. Thus assuming that subjects in general predicted the sentences to match the pictures, they should have been

able to anticipate the presence of a collective reading. Some of this activity could, however, also be associated with general anticipation for a second adjective, since intersectives exhibited a weak trend in the same direction.

Importantly, our analysis of the pre-target region in the conjunctive conditions revealed no reliable differences at the first conjunct or during the processing of *and*, despite the fact that at the presentation of the first adjective within the matching trials, it was in principle possible for the subject to predict whether the conjunction would be intersective (always the case if the first adjective described all displayed objects) or collective (always the case if the first adjective did *not* describe all displayed objects). This of course crucially assumes that subjects' predictive processing is based on the anticipation of a match. The fact that our ATL effects only emerged once the conceptual content of the second adjective was revealed suggests that either subjects did not predict as much as was possible or that these effects reflect rather bottom-up processes of conceptual combination where the actual conceptual content of the composing items matters. This latter interpretation is robustly consistent with prior findings on the LATL combinatory activity measured here, which is systemically observed even when composition is completely predictable (Bemis & Pykkänen, 2011, 2013a) and is highly sensitive to the conceptual specificity level of the composing items (Westerlund & Pykkänen, 2014; Zhang & Pykkänen, 2015).

CONCLUSION

In this work we aimed to characterize the nature of the conceptual combination that the LATL is responsible for by comparing two types of conjunctions that differed in terms of the way they distribute two features to members of a set: either the two features are added to one representation (intersective conjunction) or they are each distributed to different members of a set (collective conjunction). By showing increased activity in the LATL for intersective conjunctions over collective conjunctions, we provide evidence in favor of the hypothesis that the LATL is responsible for the combination of features to form one coherent entity representation. The pattern observed in the RATL was the opposite, indicating that its function in terms of conceptual combination is clearly distinct from the LATL.

ACKNOWLEDGEMENT

This research was supported by the National Science Foundation Grant BCS-1221723 (L.P.) and grant G1001 from the NYUAD Institute, New York University Abu Dhabi (L.P.). The first author also gratefully acknowledges the NWO (The Netherlands Organisation for Scientific Research) VICI grant number 227-80-002, “Between Logic and Common Sense: the Formal Semantics of Words” (2010-15), which allowed her to visit the NYU Neurolinguistics Lab.

REFERENCES

- Adachi, Y., Shimogawara, M., Higuchi, M., Haruta, Y., & Ochiai, M. (2001). Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. *IEEE Trans. Appl. Supercond.* 11, 669-672.
- Baron, S. G., & Osherson, D. (2011). Evidence for conceptual combination in the left anterior temporal lobe. *Neuroimage* 55, 1847-1852.
- Baron, S. G., Thompson-Schill, S. L., Weber, M., & Osherson, D. (2010). An early stage of conceptual combination: Superimposition of constituent concepts in left anterolateral temporal lobe. *Cognitive neuroscience*, 1(1), 44-51.
- Beeman, M. J., & Chiarello, C. (1998). Complementary right- and left-hemisphere language comprehension. *Current Directions in Psychological Science* 7(1), 2-8.
- Bemis, D. K., & Pykkänen, L. (2011). Simple composition: a magnetoencephalography investigation into the comprehension of minimal linguistic phrases. *Journal of Neuroscience* 31(8), 2801-2814.
- Bemis, D. K., & Pykkänen, L. (2013a). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex* 23(8), 1859-1873.
- Bemis, D. K., & Pykkänen, L. (2013b). Flexible composition: MEG evidence for the deployment of basic combinatorial linguistic mechanisms in response to task demands. *PLoS ONE* 8(9), e73949.
- Blanco-Elorrieta, E., & Pykkänen, L. (2016). Composition of complex numbers: Delineating the computational role of the left anterior temporal lobe. *NeuroImage*, 124, 194-203.
- Bornkessel-Schlesewsky, I., & Schlewsky, M. (2013). Reconciling time, space and function: a new dorsal-ventral stream model of sentence comprehension. *Brain and language*, 125(1), 60-76.
- Bornkessel-Schlesewsky, I., Schlewsky, M., Small, S. L., & Rauschecker, J. P. (2015). Neurobiological roots of language in primate audition: common computational properties. *Trends in cognitive sciences*, 19(3), 142-150.
- Brennan, J., & Pykkänen, L. (2012). The time-course and spatial distribution of brain activity associated with sentence processing. *Neuroimage* 60, 1139-1148.

- Caramazza, A., Gordon, J., Zurif, E.B., & DeLuca, D. (1976). Right-hemispheric damage and verbal problem solving behavior. *Brain and Language* 3, 41-46.
- Coutanche, M. N., & Thompson-Schill, S. L. (2014). Creating concepts from converging features in human cortex. *Cerebral Cortex*, bhu057.
- Davidson, D. (1967). The logical form of action sentences. In N. Rescher (ed.) *The logic of decision and action*. Pittsburgh: University of Pittsburgh Press. 81-94. Reprinted in Davidson (1980), 105-122.
- Del Prato, P., and Pylkkänen, L. (2014). MEG evidence for conceptual combination but not numeral quantification in the left anterior temporal lobe during language production. *Frontiers in Psychology* 5 (524).
- Done, D. J., & Gale, T. M. (1997). Attribute verification in dementia of Alzheimer type: Evidence for the preservation of distributed concept knowledge. *Cognitive Neuropsychology* 14(4), 547-571.
- Ferstl, E. C., Neumann, J., Bogler, C., & Von Cramon, D. Y. (2008). The extended language network: a meta- analysis of neuroimaging studies on text comprehension. *Human brain mapping*, 29(5), 581-593.
- 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 and Language* 74(2), 289-300.
- 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.
- Gardner, H., Silverman, J., Wapner, W., & Zurif, E. (1978). The appreciation of antonymic contrasts in aphasia. *Brain and Language* 6, 301-317.
- Gorno-Tempini, M.L., Price, C.J., 2001. Identification of famous faces and buildings: a functional neuroimaging study of semantically unique items. *Brain* 124 (10), 2087–2097.
- Grabowski, T.J., Damasio, H., Tranel, D., Boles Ponto, L.L., 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.
- Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar* (Vol. 13). Oxford: Blackwell.
- Heycock, C., & Zamparelli, R. (1999). Towards a unified analysis of DP conjunction. In *Proceedings of the Twelfth Amsterdam Colloquium*, 127-132. Universiteit van Amsterdam.
- Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: Implications for the organization of semantic memory. *Semantic Knowledge and Semantic Representations* 3(3-4), 463-495.
- Humphries, C., Binder, J. R., Medler, D. A., & Liebenthal, E. (2006). Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *Journal of Cognitive Neuroscience* 18(4), 665-679.
- Humphries, C., Love, T., Swinney, D., & Hickok, G. (2005). Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping* 26(2), 128-138.

- Humphries, C., Willard, K., Buchsbaum, B., & Hickok, G. (2001). Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *Neuroreport*, 12(8), 1749-1752.
- Jobard, G., Vigneau, M., Mayozer, B., & Tzourio-Mayozer, N. (2007). Impact of modality and linguistic complexity during reading and listening tasks. *Neuroimage* 34, 784-800.
- Keenan, E., & Faltz, L. (1985). *Boolean Semantics for Natural Language*. D. Reidel, Dordrecht.
- Krifka, M. (1990). Boolean and non-boolean 'and'. In L. Kálmán and L. Pólós (Eds.), *Papers from the Second Symposium of Logic and Language*. Akadémiai Kiadó, Budapest.
- Lambon Ralph, M. A., Cipolotti, L., Manes, F., & Patterson, K. (2010). Taking both sides: Do unilateral anterior temporal lobe lesions disrupt semantic memory? *Brain* 133(11), 3243-3255.
- 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 (4), 832.
- Lancaster, J., Rainey, L., Summerlin, J., Freitas, C., Fox, P., Evans, A., et al. (1997). Automated labeling of the human brain: a preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping* 5(4), 238
- Lancaster, J., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping* 10(3), 120-131.
- Leffell, T., Lauter, M., Westerlund, M., & Pylkkänen, L. (2014). Restrictive vs. non-restrictive composition: a magnetoencephalography study. *Language, cognition and neuroscience*, 29(10), 1191-1204.
- Mayozer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., et al. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience* 5(4), 467-479.
- Montague, R. (1970). Universal grammar. *Theoria* 36, 373-98.
- Pallier, C., Devauchelle, A. D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences* 108(6), 2522-2527.
- Parsons, T. (1990). *Events in the Semantics of English* (Vol. 5). Cambridge, Ma: MIT Press.
- Partee, B., & Rooth, M. (1983). Generalized conjunction and type ambiguity. In R. Bauerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, Use and Interpretation of Language*. Berlin: De Gruyter.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience* 8(12), 976-987.
- Poortman, E. B. (2014). Between intersective and 'split' interpretations of predicate conjunction: The role of typicality. In J. Degen, M. Franke, & N. Goodman (Eds.), *Proceedings of Formal & Experimental Pragmatics 2014*, 36-42. Tübingen.
- Pylkkänen, L., Bemis, D. K., & Elorrieta, E. B. (2014). Building phrases in language production: An MEG study of simple composition. *Cognition*, 133(2), 371-384.

- Rice, G. E., Ralph, M. A. L., & Hoffman, P. (2015). The Roles of Left Versus Right Anterior Temporal Lobes in Conceptual Knowledge: An ALE Meta-analysis of 97 Functional Neuroimaging Studies. *Cerebral Cortex*, bhv024.
- Rogalsky, C., & Hickok, G. (2009). Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cerebral Cortex* 19(4), 786-796.
- Rogers, T. T., Hocking, J., Noppeney, U., Mechelli, A., Gorno-Tempini, M. L., Patterson, K., et al. (2006). Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective, & Behavioral Neuroscience* 6(3), 201-213.
- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review* 111(1), 205
- 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.
- Vandenberghe, R., Nobre, A. C., & Price, C. J. (2002). The response of left temporal cortex to sentences. *Journal of Cognitive Neuroscience* 14(4), 550-560.
- Visser, M., Jefferies, E., & Ralph, M. L. (2010). Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, 22(6), 1083-1094.
- Warrington, E. K. (1975). The selective impairment of semantic memory. *The quarterly journal of experimental psychology* 27(4), 635-657.
- Westerlund, M., and Pykkänen, L. (2014). The role of the left anterior temporal lobe in semantic composition vs. semantic memory. *Neuropsychologia* 57, 59-70.
- Winter, Y. (2001). Plural predication and the strongest meaning hypothesis. *Journal of Semantics* 18, 333-365.
- Wong, C., & Gallate, J. (2012). The function of the anterior temporal lobe: a review of the empirical evidence. *Brain research*, 1449, 94-116.
- Zhang, L., & Pykkänen, L. (2015). The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. *NeuroImage*, 111, 228-240.

FIGURE LEGENDS

FIGURE 1: Experimental design and trial structure.

FIGURE 2: ROI results for left and right BA 38 during the processing of the sentence final adjective. The dark shading indicates significant clusters ($p < .05$ corrected) and the light shading clusters that did not survive correction for multiple comparisons. In the bottom, liberally thresholded uncorrected full brain contrasts for the same pair-wise comparisons (see Methods), calculated in order to visualize the spatial extent of the effects obtained in the ROI analysis (in yellow circles). The effects are centered on the temporal pole, with potential spread to adjacent dorsal and ventral (though not lateral) cortex.

FIGURE 3: Time course of left and right BA 38 activity during the entire conjunction. Activity during the presentation of the first adjective and the conjunctive *and* is plotted on the left and activity elicited by the second, target adjective is plotted on the right. The analysis of the first adjective and the conjunctive *and* revealed no reliable effects, suggesting that the ATL relevant computational distinction between the two conditions only takes place at the second adjective. The trial structure at the bottom is temporally aligned with the waveform data, with zero's on the waveform x-axes representing the onsets of the two adjectives.

FIGURE 1

FIGURE 2

FIGURE 3