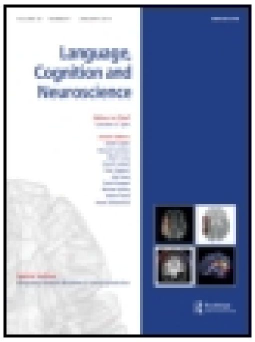

Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: http://www.tandfonline.com/loi/plcp21

Linguistic structure across time: ERP responses to coordinated and uncoordinated noun phrases

Ellen Lau & Chia-Hsuan Liao

To cite this article: Ellen Lau & Chia-Hsuan Liao (2017): Linguistic structure across time: ERP responses to coordinated and uncoordinated noun phrases, Language, Cognition and Neuroscience, DOI: 10.1080/23273798.2017.1400081

To link to this article: http://dx.doi.org/10.1080/23273798.2017.1400081

LANGUAGE, COGNITION AND NEUROSCIENCE, 2017
https://doi.org/10.1080/23273798.2017.1400081

REGULAR ARTICLE

Linguistic structure across time: ERP responses to coordinated and uncoordinated noun phrases

Ellen Lau and Chia-Hsuan Liao

Department of Linguistics, Neuroscience and Cognitive Science Program, Maryland Language Science Center, University of Maryland, College Park, MD, USA

ABSTRACT
Relatively little is known about how linguistic structure is neurally encoded. The current study examines a relatively subtle manipulation of syntactic and semantic structure: the difference between reading a list of two noun phrases ("sunlit ponds ### green umbrellas") and their syntactic coordination ("sunlit ponds and green umbrellas"). In two ERP experiments, the presence of the coordinator resulted in an increased anterior negativity across the entire second noun phrase, even though coordination had no direct relevance for the memory recognition task. These findings demonstrate that structural connectedness exerts strong, ongoing differences in neural activity even when structured and unstructured materials are very tightly matched in sequence and content. These differences may reflect ongoing maintenance of structure in memory, or computation of the more complex semantic or discourse representation associated with syntactic coordination.

ARTICLE HISTORY
Received 22 April 2017
Accepted 23 October 2017

KEYWORDS
Syntax; event-related potentials; parsing; coordination

Introduction

Accurate language comprehension requires encoding syntactic relationships between the words (or morphemes) of a sentence, using knowledge of the grammatical rules of the language. For example, in representing the 7-word phrase *The red bowl that my mother loves …* , encoding of both local (*mother – loves*) and non-local (*loves – bowl*) relationships is required. Many of the broad generalisations governing these relations are well understood. We also know from previous neuroimaging work that structured sentences elicit increased neural responses compared to unstructured word lists in a number of brain regions, including left inferior frontal gyrus and left anterior and posterior temporal cortex (Bemis & Pylkkänen, 2011; Fedorenko, Hsieh, Nieto-Castañon, Whitfield-Gabrieli, & Kanwisher, 2010; Friederici, Meyer, & Von Cramon, 2000; Humphries, Binder, Medler, & Liebenthal, 2006; Mazoyer et al., 1993; Pallier, Devauchelle, & Dehaene, 2011; Rogalsky & Hickok, 2009; Stowe et al., 1998; Vandenberghe, Nobre, & Price, 2002). However, less is known about how structured linguistic information is neurally represented across time, during the short-term window in which a sentence is parsed and interpreted.

On the basis of their fMRI results, Pallier et al. (2011) have recently proposed that linguistic structure is neurally encoded in such a way that overall neural activity will systematically increase across the timecourse of a structurally connected constituent until that constituent is completed. The current ERP study tests predictions of their proposal with a novel comparison that manipulates structural connectedness while tightly controlling the sequence and lexical content of the critical stimuli and introducing relatively minimal differences in semantic content: the syntactic coordination of two noun phrases (*sunlit ponds and green umbrellas*) vs. lists of two noun phrases separated by a placeholder symbol (*sunlit ponds ### green umbrellas*). If prior linguistic structure is represented across the course of a constituent in a way that causes neural activity to increase with the size of the connected structures, sustained differences in the ERP are predicted across the course of the second noun phrase when it is syntactically connected to the prior material vs. when it is not.

Background

Syntactic knowledge provides instructions for combining the meanings associated with phonological forms, to ultimately feed computation of the message intended by the speaker. Surface forms of different phonological length and structural complexity can be associated

CONTACT Ellen Lau ellenlau@umd.edu

Supplemental data for this article can be accessed https://doi.org/10.1080/23273798.2017.1400081.

with a similar meaning (e.g. "yellow buses" and "buses that are yellow"), and prior work has indicated that after a sentence is over, memory for its message persists over long intervals but memory for its form is relatively fragile and short-lived (e.g. Sachs, 1967). However, during the course of the sentence, aspects of the structural form must persist in "short-term" memory (STM) in order to make non-adjacent structural relations possible.

In a seminal fMRI study, Pallier et al. (2011) propose and test a specific hypothesis about how this short-term representation of sentence structure is neurally encoded. Pallier et al. (2011) begin by adopting Smolensky and Legendre's (2006) model of syntactic composition. In this model, syntactic structure is represented by superposition of a number of tensor products of "role" and "filler" vectors, where each node in the tree structure is represented by the tensor product of the role vector representing the position in the tree and the filler vector representing the content (e.g. lexical entry) in that position. Pallier et al. then assume that the filler vector will have a sparse neural code such that each new filler will recruit a largely new set of neurons, and they assume that (at least for right-branching structures) the full tree has to be actively maintained across the course of the constituent being built. They also assume that (at least for right-branching structures) these products are *actively maintained* across the span of a connected syntactic structure such that the longer the structure, the larger the number of active neural units. This yields the prediction that for sentences or phrases with right-branching structures, neural activity will monotonically increase across the course of the phrase, resetting only when the phrase is completed.

In order to test their hypothesis, Pallier et al. (2011) used a novel *constituency-size* manipulation. This design holds constant the number of words in a trial (12) while varying the size of the connected structure chunks that they compose (an unstructured list of 12 words, a sequence of 6 unconnected two-word phrases, a sequence of 4 three-word phrases, a sequence of 3 four-word phrases, a sequence of 2 six-word phrases, and a twelve-word right-branching sentence). Although the temporal resolution of the BOLD (blood-oxygenation-level dependent) signal in fMRI is too poor to precisely resolve the increase in activity at each incoming word predicted by the model, the integrated BOLD signal across the entire trial should increase parametrically with the length of the longest constituent. As predicted by their active structure maintenance model, Pallier et al. found a number of left-hemisphere regions whose activity increased monotonically with the number of lexical-structure bindings that needed to be maintained as part of a connected structure across time: inferior frontal gyrus (IFG), anterior and posterior temporal cortex (pSTS), and angular gyrus. Pallier et al. also included a jabberwocky manipulation, in which the content words were replaced by meaningless pseudowords, in order to distinguish active structure maintenance effects from processes associated with conceptual combination, world knowledge, and discourse model differences in smaller and larger phrase structures. They found that IFG and pSTS continued to show constituency-size effects for these jabberwocky materials, and took this as evidence that these regions indeed supported active maintenance of sentence structure.

Although Pallier et al.'s (2011) findings are suggestive, they are not conclusive. First, while the parametric design across a number of levels of constituent size has obvious virtues, it prevents certain kinds of stimulus control that could leave open alternative explanations. For example, the paradigm does not control lexical content across conditions, such that the proportion of function words varied across constituency size. These lexical differences, rather than structure maintenance, could have been responsible for some of the differences they observed. The randomised presentation they used also meant that participants did not know the extent to which an upcoming trial would be structured or not, so that some of the effects observed might have resulted from the degree to which a given trial was expected to be a full sentence rather than structure maintenance per se. A recent follow-up study addressed several of these concerns in a partial replication of the constituency-size manipulation using lexically-controlled materials in the 1-word, 2-word, and 6-word conditions (Matchin, Hammerly, & Lau, 2017), and notably failed to find differences between the 1-word (list) condition and the 2-word phrase condition as their model would have predicted. However, Matchin et al. (2017) did replicate their finding of increased activity for sentences vs. lists in IFG and pSTS (among other regions), even though the sentences were simple, contained only local dependencies, and were completely uniform in structure across the experiment, and this result is consistent with active maintenance of previous or predicted structure.

The poor temporal resolution of fMRI also makes it non-ideal for testing one of the critical predictions of Pallier et al.'s proposal, which is that activation will increase across the course of the connected structure. EEG, MEG, or intracranial recordings can better test this prediction, and there are some interesting suggestions of this in the existing literature. Van Petten and Kutas (1991) reported a slow positive shift across the course of scrambled word lists compared to sentences in EEG, which they attributed to differences in arousal but which might also be consistent with an increased negative shift associated with structure maintenance. Bastiaansen, Magyari, and Hagoort (2010) observed increasing activity in theta and beta power across the course of sentences compared to scrambled word lists in EEG, and attribute the theta increase to a "working memory trace" of the input. Fedorenko et al. (2016) compare scrambled word lists and sentences in electrocorticography (ECoG) patients and find a monotonic increase in gamma power for regular sentences but not for jabberwocky sentences, and therefore attribute the increase to the representation of complex meanings. Although this prior work is suggestive, contrasting sentences with scrambled sentences is somewhat messy in temporally sensitive techniques because the distribution of word classes within the trial differs across conditions; for example, in English a verb is more likely to occur in the first position of a scrambled sentence than in a sentence. Recent work by Ding and colleagues (Ding, Melloni, Zhang, Tian, & Poeppel, 2015) argues that neural representation of structure across the course of the sentence can be observed with a very different constituent *rate* approach. Ding et al. recorded MEG while presenting syllables at a constant rate, organised such that they also formed phrase and sentence-level constituents at a constant rate (e.g. Adj-N-V-N-Adj-N-V-N-Adj-N-V-N …). Ding et al. observed peaks in MEG spectral power not only at the syllable rate frequency, but also at the phrasal rate and the sentence rate when the sentences were parsable by the listeners (e.g. for native speakers of the language and not for non-speakers of the language), and suggest that the increased power of the MEG response at these frequencies may index "the neural representation of the internally constructed hierarchical linguistic structure". However, although these data provide strong neural evidence for online syntactic structure-building, it is not clear that they speak to Pallier et al.'s proposal that maintaining the syntactic representation requires sustained neural activity across the course of a constituent, because the pattern Ding et al. observe could easily reflect punctate structure-building processes that are triggered at critical points in the sentence, rather than sustained activity across time (oscillatory or otherwise).

Finally, it is worth noting that Pallier et al.'s (2011) active maintenance structure encoding proposal makes assumptions about the memory architecture that are a matter of vigorous debate in the sentence processing literature. On the one hand, as classically reviewed in Caplan and Waters (1999), a long tradition of work in sentence processing has assumed, like Pallier et al., that sentence structure is represented over the short term in a qualitatively distinct state from long-term memory. In particular, these traditional accounts assume that memory for sentence structure is supported by a working memory buffer and argue that many cases of sentence processing difficulty could be attributed to exceeding the capacity limit of this buffer (e.g. Gibson, 1998; Kimball, 1973; Miller & Isard, 1964; Yngve, 1960). On the other hand, in recent years several groups have argued that sentence processing is better modelled as a series of retrievals of sentence subparts from long-term memory, and that processing difficulty can instead be explained through retrieval interference, so that no active maintenance of structure need be assumed (Lewis & Vasishth, 2005; McElree, Foraker, & Dyer, 2003; Van Dyke & Johns, 2012). Given the assumptions of the latter accounts, the structure of the current sentence could exist only in the abstract as a collection of long-term memory chunks that are never all active at the same time, and there would be *no* sense in which we could expect to see reflections of this kind of representation itself in the patterns of broadly distributed neural activity recorded by non-invasive cognitive neuroscience techniques such as EEG, MEG, and fMRI. Rather, we could only expect to see indirect reflections of the long-term memory retrieval process, such as increased activity when competition between multiple interfering active representations needs to be resolved (e.g. Glaser, Martin, Van Dyke, Hamilton, & Tan 2013). This mirrors a longstanding and as yet unresolved debate in cognitive neuroscience about whether or not short-term or working memory representations are qualitatively distinct from long-term memory in being neurally encoded through sustained activity (e.g. Funahashi, Bruce, & Goldman-Rakic, 1989; Fuster & Alexander, 1971; Lewis-Peacock, Drysdale, Oberauer, & Postle, 2012; Lundqvist et al., 2016; Postle, 2006; Stokes, 2015).

The current study

Given the methodological and theoretical considerations discussed above, it seems clear that further evaluation of Pallier et al.'s (2011) structure encoding hypothesis is needed. In the current study we use a more time-sensitive measure (event-related potentials) and we use phrasal structures rather than full sentences to allow for a better controlled design: a novel list/structure paradigm that compares noun phrase coordination to unconnected noun phrase list sequences (Figure 1). In unconnected list trials, two noun phrases are presented word-by-word, separated by a placeholder symbol (###). In coordination trials, the two noun phrases are coordinated with *and*. After occasional trials, participants

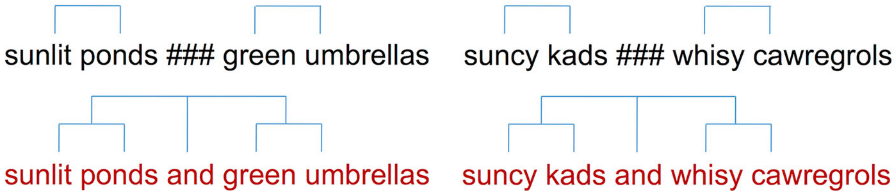

Figure 1. Illustration of the design, contrasting lists of two noun phrases with structurally connected coordinate phrases.

are given a noun phrase recognition memory probe to encourage attention. The prediction of the active structure maintenance hypothesis is fairly straightforward; in the coordination condition, the second noun phrase is structurally connected to the first, and therefore the terminals of the first noun phrase need to be actively maintained, but this is not the case in the list condition. Therefore, this hypothesis predicts a sustained difference in activity across the second noun phrase (beyond any short-lived differences associated with processing the coordinator vs. the placeholder). Any non-structural rehearsal strategies for performing the memory task should be constant across conditions.

This design has a number of virtues for testing the active maintenance hypothesis. First, the critical region – the second noun phrase – is exactly matched in physical properties and position within the trial across conditions. Second, any differences in meaning between the coordination trials and the unconnected noun phrase trials are fairly subtle. Third, encountering unconnected noun phrases is not as unusual or disruptive as encountering scrambled sentences and seems less likely to drive differences in attentional state or task strategy. Finally, this design isolates the maintenance of prior structure rather than predicted structure, as the predicted categories at the critical measurement period across the second noun phrase are the same (e.g. the prediction of a noun during the presentation of the adjective).

We included both natural (real word) and jabberwocky conditions. As discussed above, Pallier et al. (2011) were particularly interested in demonstrating evidence for the maintenance of *syntactic* structure, and argued that observing constituent-size effects for jabberwocky materials as well as natural materials constitutes a strong argument that the effects are syntactic in nature. Although there are some reasons to question the strength of this argument (many sentence-level semantic relations such as predicate-argument relations could be computed on a syntactically well-formed string without access to lexical semantics, and jabberwocky materials may engender less attention and therefore more shallow structural analysis than natural ones), we included this manipulation for the purpose of better comparing the results across the two designs.

Experiment 1

Methods

Materials

The experiment comprised a 2 × 2 design (structure: coordination/list × content: natural/jabberwocky). In total, 50 trials were presented in each of the 4 conditions, with each trial containing two adjective–noun sequences.

200 different adjectives and 200 different nouns were selected and combined to form 200 natural adjective–noun pairs that each formed a meaningful and plausible phrase. In order to minimise prediction of specific lexical items from the adjective to the noun, all adjectives were "low-constraint"; that is, in the Corpus of Contemporary American English, no single word followed a given adjective more than 15% of the time. The noun was always presented in the plural form so that the phrase could be construed as a complete (bare) plural determiner phrase in the absence of an article. Many of the adjectives contained a common adjectival affix (e.g. -y, -al, -able, -ful, -ous, -ic). The 200 adjective–noun phrases were then randomly paired into 100 sets of two. These 100 item sets were then distributed across four lists in a Latin Square design, such that each item would appear in the coordination or no-coordination condition, and such that the phrases could appear in either order (e.g. *new books ### shiny floors; shiny floors ### new books; new books and shiny floors; shiny floors and new books*).

200 jabberwocky "adjectives" were created by combining a pronounceable nonword stem (e.g. *blarg*) with a common adjectival affix (e.g. *blargful*). 200 jabberwocky "nouns" were created by combining a non-overlapping set of pronounceable nonword stems (e.g. *kropt*) with the plural –*s* affix (e.g. *kropts*). The nonwords were derived using the Wuggy package in order to match lexical and sublexical properties of the real words (Keuleers & Brysbaert, 2010). The 200 jabberwocky adjective–noun phrases were then randomly paired into 100 sets of two. These 100 item sets were then distributed across the four lists in a Latin Square design, such that each item would appear in the coordination or no-coordination condition, and such that the phrases could appear in either order.

To instantiate the memory task, we added memory probes to 10 items in each condition, for a total of 40 memory probes across the experiment (20% of trials). Half of the probes were identical to one of the previous items, requiring a "YES" response, and half were not, requiring a "NO" response. The "NO" items were sometimes incorrect combinations of words from the previous trial (e.g. given a trial like *new books and shiny floors* the probe might be *new floors*) and sometimes combined one of the previous words with a new word (e.g. *polished floors*).

Procedure

In Experiment 1, we chose to block the structure manipulation, such that the first half of the experiment included only list trials and the second half of the experiment included only coordinate trials. We did this because we were concerned that participants would be likely to mentally insert a linguistic coordinator in the list condition if given the idea to do so by seeing the coordination conditions (note that strictly speaking, an overt coordinator is not required for linguistic coordination in English, e.g. *I see birds, flowers, trees …*). This would result in equal constituent size across conditions. To prevent this, in Experiment 1 the experimental materials were divided into two presentation blocks, the first of which included 100 randomly distributed natural and jabberwocky list trials, and the second of which included 100 randomly distributed natural and jabberwocky coordinated trials. A break screen appeared halfway through each block.

A practice session preceded the experimental session in which participants were presented with a number of list trials. They were told that each trial would contain two phrases and that after some trials they would be presented with a probe phrase and asked to indicate whether they had seen this phrase or not. In between the two blocks, participants were told that now "the two phrases would be connected, but that the task would remain the same" and were given several practice coordination trials to illustrate.

During the experiment, participants were seated in a chair in a dimly lit room. Stimuli were visually presented one word at a time on a computer monitor in white 24-point case Arial font on a black background. Each trial began with a 1500 ms fixation cross. After a 200 ms blank screen, the five stimuli of each trial were presented at a rate of 500 ms on and 100 ms off. The final word stayed on the screen for 700 ms, followed by a blank screen of 300 ms. After 20% of the trials, the blank screen was followed by a memory probe (e.g. *suncy kads?*) which remained on the screen until the participant responded.

Participants

A total of 41 students (23 females, 18–31 years old, mean age: 20) from University of Maryland participated in the current study. All the participants were native speakers of English, and none of them reported a history of neurological or psychiatric disorder or recent use of psychoactive medications. Since the duration of our critical epoch was quite long (i.e. 2000 ms), after data pre-processing, 12 participants had to be excluded from further ERP analysis due to excessive eye-blinking, muscle potentials and alpha waves during the epoch. Muscle potential and alpha wave artifacts were identified using the peak-to-peak artifact rejection routine provided by ERPLAB, and eye-blink and eye-movement artifacts were identified using the step function artifact rejection routine provided by ERPLAB. Participants were excluded only when more than 40% of trials were rejected. The reported results were obtained from the remaining 29 participants (16 females, 18–23 years old, mean age: 20). All the participants were right-handed as assessed by the Edinburg Handedness Inventory (Oldfield, 1971). Written informed consent was obtained from all participants.

Electrophysiological recording

Twenty-nine tin electrodes were held in place on the scalp by an elastic cap (Electro-Cap International, Inc., Eaton, OH) in a 10–20 configuration (O1, O2, P7, P3, Pz, P4, P8, TP7, Cp3, CPz, CP4, TP8, T7, C3, Cz, C4, T8, FT7, FC3, FCz, FC4, FT8, F7, F3, Fz, F4, F8, FP1, FP2). Bipolar electrodes were placed above and below the left eye and at the outer canthus of the right and left eyes to monitor vertical and horizontal eye movements. Additional electrodes were placed over the left and right mastoids. Scalp electrodes were referenced online to the left mastoid and re-referenced offline to the average of left and right mastoids. The ground electrode was positioned on the scalp between Fz and Cz. Impedances were maintained at less than 10 kΩ for all scalp and ocular electrode sites and less than 2 kΩ for mastoid sites. The EEG signal was amplified by a NeuroScan SynAmps® Model 5083 (NeuroScan, Inc., Charlotte, NC) with a bandpass of 0.05–100 Hz and was continuously sampled at 500 Hz by an analog-to-digital converter.

Analysis

We extracted epochs time-locked to the onset of the coordinator or corresponding place-holder from −100:2000 ms, thus extending from the coordinator through the response to the second noun phrase. Averaged ERPs were formed off-line from these epochs, after rejecting trials free of ocular and muscular artifact, using preprocessing routines from the EEGLAB (Delorme & Makeig, 2004) and ERPLAB (Lopez-Calderon & Luck, 2014) toolboxes. Across the 29 participants included in the analysis, approximately 21% of the trials were rejected because of artifact. The exact rejection rates by condition for the critical epochs (time-locked to the coordinator) were: list-natural – 17%, list-jabberwocky – 21%, coordination-natural – 18%, coordination-jabberwocky – 19%. A 100-ms pre-stimulus baseline was subtracted from all waveforms, and a 40-Hz low-pass filter was applied to the ERPs offline. ERP data for this and the other experiment in this paper are publicly available on the first author's website (http://ling.umd.edu/~ellenlau/public_data_archive/ERP_coordination/Lau&Liao_2017).

Because we were interested in evaluating the presence of sustained activity associated with structure maintenance, we conducted analyses on mean ERP amplitudes in a broad 100–600 ms time-window after the onset of each of the three words in the epoch: the coordinator, 100–600 ms, the second adjective, 700–1200 ms, and the second noun, 1300–1800 ms (we did not include the first 100 ms post-onset in the analysis as we assumed that at least this much time would be needed for basic sensory processing). In order to quantify the topography of the effects observed, we focused on a subset of 16 electrodes (left anterior: F7, F3, FT7, FC3; right anterior: F4, F8, FC4, FT8; left posterior: TP7, CP3, P7, P3; right posterior: CP4, TP8, P4, P8) and used R (R Development Core Team, 2013) to conduct a quadrant analysis consisting of a 2 × 2 × 2 × 2 (structure × content × hemisphere × anteriority) Type III SS repeated-measures ANOVA. We report all significant main effects and interactions involving the factors of structure and content. Figures in text present ERP waveforms at selected electrodes; full ERP data from all scalp electrodes are included as supplementary figures.

Table 1. *P*-values corresponding to comparisons that included the factors of coordination or content in the omnibus ANOVAs conducted for ERPs in the three time-windows of interest in Experiment 1.

	Coordinator (100–600 ms)	Adj. 2 (700–1200 ms)	Noun 2 (1300–1800 ms)
Coordination	**$p < .001$**	**$p = .009$**	$p = .08$
Content	**$p = .03$**	$p = .12$	$p = .2$
Coordination × Content	$p = .11$	$p = .08$	$p = .1$
Coordination × Hemisphere	$p = .2$	$p = .07$	$p = .2$
Coordination × Anteriority	**$p = .003$**	**$p = .01$**	**$p = .02$**
Coordination × Hemisphere × Anteriority	$p = .2$	$p = .09$	$p = .3$
Content × Hemisphere	$p = .5$	**$p = .01$**	$p = .3$
Content × Anteriority	$p = .4$	$p = .3$	$p = .3$
Content × Hemisphere × Anteriority	**$p = .03$**	$p = .1$	$p = .07$
Coordination × Content × Hemisphere	$p = .3$	$p = .9$	$p = .3$
Coordination × Content × Anteriority	$p = .9$	$p = .97$	$p = .2$
Coordination × Content × Hemisphere × Anteriority	$p = .9$	$p = .95$	$p = .96$

Note: Values below .05 are indicated in bold.

As we used a block design in Experiment 1, it is possible that responses also differed prior to the coordinator or corresponding placeholder across blocks. In order to evaluate this possibility we also computed epochs time-locked to the onset of the first adjective (at the beginning of the trial) and extending through the first noun phrase (−100 ms:1200 ms, and baselined these epochs to the 100 ms preceding the onset of the first noun phrase.

Results

Behavioural

Overall accuracy on the memory task was 85.4% (natural list: 91.0%; natural coordination: 92.0%; jabberwocky list: 80.0%; jabberwocky coordination: 78.6%). A 2 × 2 repeated measure ANOVA revealed a main effect of content ($F(1, 28) = 37.1$, $\mathrm{MSE} = 4346$, $p < .01$) reflecting lower accuracy in the jabberwocky conditions. There was no main effect of structure ($F(1, 28) = .006$, $\mathrm{MSE} = .86$, $p > .2$), and no interaction between structure and content ($F(1,28) = .476$, $\mathrm{MSE} = 42.24$, $p > .2$).

ERPs

Visual inspection of the grand-average ERPs time-locked to the coordinator position indicates that responses were more negative throughout the epoch for the coordination condition relative to the non-coordination condition. Topographically, the distribution of this effect shifted across the epoch: at the coordinator position, coordination resulted in an increased posterior negativity in both natural and jabberwocky conditions, while during the subsequent noun phrase, coordination resulted in a sustained increased anterior negativity that was mainly present in the natural conditions (Figures 2 and 3).

Statistical analysis confirmed these observations (Table 1). In the 100–600 ms time-window corresponding to 100 ms post-onset of the coordinator position, an omnibus quadrant ANOVA revealed a significant main effect of coordination ($F(1,28) = 18.7$, $\mathrm{MSE} = 6.4$, $p < .01$) and a significant interaction between coordination and anteriority ($F(1,28) = 10.3$, $\mathrm{MSE} = .61$, $p < .01$), as well as a main effect of content ($F(1,28) = 5.3$, $\mathrm{MSE} = 2.94$, $p < .05$) and an interaction between content, hemisphere, and anteriority ($F(1,28) = 5.2$, $\mathrm{MSE} = .11$, $p < .05$). The interaction between coordination and anteriority was driven by a larger coordination effect in posterior channels (1.2 μV) than in anterior channels (.8 μV). These effects are likely to reflect differential early visual processes and lexical access and grammatical analysis of

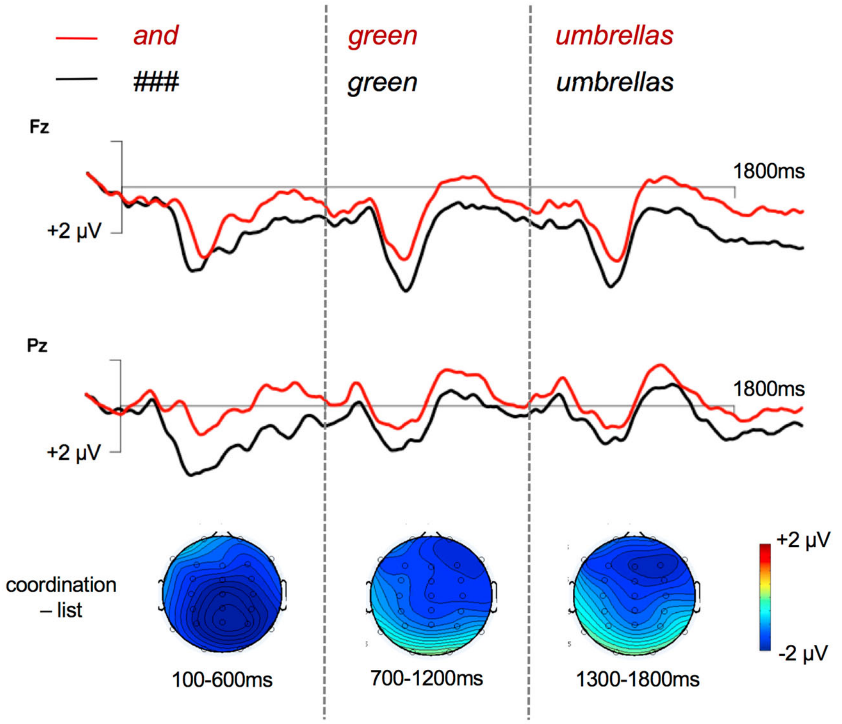

Figure 2. Scalp maps and event-related potentials time-locked to the coordinator (*and*) or the placeholder (*###*) at selected frontal and posterior electrode sites for natural list and natural coordination conditions in Experiment 1. Scalp maps illustrate the mean difference between coordination and list conditions across a 100–600 ms time-window following presentation of each word.

the coordinator (*and*) relative to the placeholder (*###*) in both natural and jabberwocky trials; the central-posterior distribution is similar to N400 effects of lexical-conceptual processing.

In the 700–1200 ms time-window corresponding to 100 ms post-onset of the adjective of the second noun phrase, we again observed a main effect of coordination ($F(1,28) = 7.9$, MSE = 8.4, $p < .05$) and a significant interaction between coordination and anteriority ($F(1,28) = 6.8$, MSE = 1.0, $p < .05$), but this time the interaction was driven by a larger coordination effect in anterior electrodes (1.0 μV) than in posterior electrodes (.5 μV). We also observed a marginally significant interaction between coordination and content ($F(1,28) = 3.3$, MSE = 5.39, $p = .08$), driven by a larger coordination effect in natural (1.1 μV) than jabberwocky (.4 μV). There was also a significant interaction between content and hemisphere ($F(1,28) = 7.5$, MSE = .85, $p < .05$), driven by more positive responses to jabberwocky in the left hemisphere and more positive responses to natural text in the right hemisphere.

The pattern was similar in the 1300–1800 ms time-window corresponding to 100 ms post-onset of the noun of the second noun phrase. Here we observed a marginal main effect of coordination ($F(1,28) = 3.4$, MSE = 8.4, $p = .08$) and again a significant interaction between coordination and anteriority ($F(1,28) = 5.6$, MSE = 1.57, $p < .05$), again driven by a larger coordination effect in anterior electrodes (.8 μV) than posterior electrodes (.2 μV). Although the effect appeared larger in English conditions, the interaction between coordination and content only trended towards significance ($F(1,28) = 2.7$, MSE = 7.5, $p = .1$).

Because Experiment 1 used a block design, such that participants knew in advance that trials in the first block were list trials and trials in the second block were coordination trials, one potential concern is that the results we observe during the second noun phrase were ongoing effects from the first noun phrase, or due to artifactual baseline shifts. Supplementary Figure 3 illustrates the ERP for the first noun phrase, which shows that responses were tightly matched prior to the onset of the coordinator or placeholder.

Discussion

In Experiment 1, we presented participants with an initial block of intermixed natural and jabberwocky trials comprising two separate noun phrases, and a subsequent block of intermixed natural and jabberwocky trials comprising two coordinated noun phrases. Unsurprisingly, we observed differential activity at the coordinator (*and*) relative to the control stimulus (*###*) presented in the list condition, which took the form of an increased

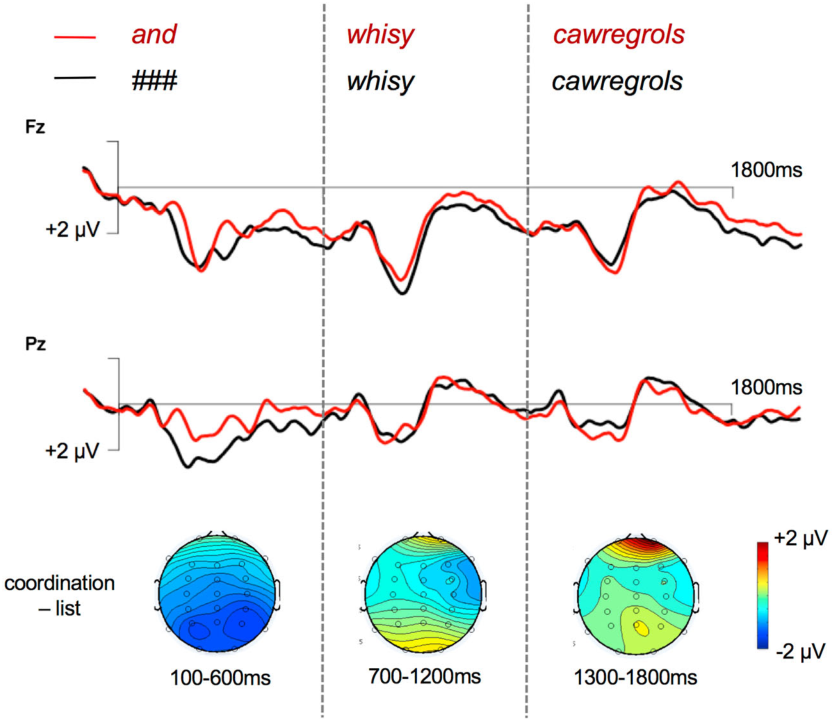

Figure 3. Scalp maps and event-related potentials time-locked to the coordinator (*and*) or the placeholder (*###*) at selected frontal and posterior electrode sites for jabberwocky list and jabberwocky coordination conditions in Experiment 1. Scalp maps illustrate the mean difference between coordination and list conditions across a 100–600 ms time-window following presentation of each word.

posterior negativity for the coordinator between ~100–600 ms. However, in the natural conditions we also observed downstream effects of coordination on the processing of the second noun phrase, which manifested as a sustained increased anterior negativity throughout the second noun phrase in the coordination condition relative to the list condition.

The results of the structural manipulation in the natural trials appear consistent with an active structure maintenance account in which prior structure must be actively maintained across the course of a connected phrase. According to such an account, the increased anterior negativity that we observed during the second noun phrase in coordination trials would reflect the activity of the neural assembly encoding the structural position of the *first* noun phrase in the coordinated NP – in other words, the neural activity required to maintain the binding of the lexical terminals in the first noun phrase to positions in a relational syntactic structure representation. Interestingly, sustained anterior negativities have been previously associated with short-term memory maintenance of phonological information (Ruchkin, Johnson, Canoune, & Ritter, 1990; Ruchkin, Johnson, Grafman, Canoune, & Ritter, 1992) and syntactic predictions (Fiebach, Schlesewsky, & Friederici, 2002; King & Kutas, 1995; Phillips, Kazanina, & Abada, 2005; Ueno & Kluender, 2003), and therefore the distribution and topography of the current structural connectedness effect might suggest a unified mechanism across these various short-term memory computations.

However, the structural connectedness effect seemed to be largely absent in the jabberwocky conditions, while Pallier et al.'s (2011) proposal holds that the cost of maintaining structure should be independent of lexical semantic content. There are at least two possible explanations for this contrast between the natural and jabberwocky results. One possibility is that the sustained anterior negativity observed in the natural conditions relates to the encoding of semantic rather than syntactic structural representations, in which case the current ERP results would constitute negative evidence against Pallier et al.'s syntactic structure encoding account (a failure to observe a predicted increase in neural activity). The other possibility is that the sustained anterior negativity we observe in the natural conditions does indeed reflect the structure maintenance activity posited by Pallier et al. (2011), but that the short adjective–noun phrases used here did not contain enough functional content to encourage participants to assign a syntactic structure to them in the jabberwocky conditions. We review these possibilities in more detail in the General Discussion.

Experiment 2 was designed to determine whether the sustained anterior negativity would also be observed in a randomised design in which coordination and list trials were intermixed. We used a block design in Experiment 1 because we were concerned that randomising might result in participants applying syntactic coordination to the list trials, in the absence of an overt coordinator. However, block designs often raise concerns that state-level differences other than the one intended, such as attention or fatigue, might be driving the effects of interest. Although it was not clear to us that such a plausible alternative account existed in the current case, if it turned out that the structural effect was robust to the randomised design then we would be able to fully rule out this possibility. Experiment 2 also provided us an opportunity to replicate the novel effect that we observed in Experiment 1.

Experiment 2

Methods

Materials

Experimental materials were identical to those used in Experiment 1.

Procedure

Experimental procedure was identical to that used in Experiment 1 except that the list and coordination trials (of both types) were randomly interleaved rather than blocked. Participants were told that they would see trials that included two noun phrases and would be asked to respond to memory probes after selected trials. The practice session preceding the experiment included both coordination and list trials.

Table 2. *P*-values corresponding to comparisons that included the factors of coordination or content in the omnibus ANOVAs conducted for ERPs in the three time-windows of interest in Experiment 2.

	Coordinator (100–600 ms)	Adj. 2 (700–1200 ms)	Noun 2 (1300–1800 ms)
Coordination	**$p = .001$**	$p = .05$	$p = .3$
Content	$p = .97$	$p = .3$	$p = .2$
Coordination × Content	$p = .08$	**$p = .03$**	**$p = .045$**
Coordination × Hemisphere	**$p = .01$**	**$p = .02$**	$p = .3$
Coordination × Ant/Post	$p = .1$	$p = .3$	$p = .09$
Coordination × Hemisphere × Ant/Post	**$p = .006$**	$p = .5$	$p = .6$
Content × Hemisphere	$p = .4$	$p = .3$	$p = .1$
Content × Ant/Post	$p = .5$	$p = .7$	$p = .3$
Content × Hemisphere × Ant/Post	$p = .06$	$p = .3$	$p = .7$
Coordination × Content × Hemisphere	$p = .4$	$p = .4$	$p = .15$
Coordination × Content × Ant/Post	$p = .4$	$p = .7$	$p = .9$
Coordination × Content × Hemisphere × Ant/Post	$p = .6$	$p = .6$	$p = .6$

Note: Values below .05 are indicated in bold.

Participants

A total of 22 students (12 females, 19–26 years of age, mean age: 21) from University of Maryland participated in the current study. All the participants were native speakers of English, and none of them reported a history of neurological or psychiatric disorder or recent use of psychoactive medications. Five participants had to be excluded from further ERP analysis due to excessive eye-blinking, muscle potentials and alpha waves during the epoch. The reported results were obtained from the remaining 17 participants (10 females, 19–26 years old, mean age: 21). All the participants were right-handed as assessed by the Edinburg Handedness Inventory (Oldfield, 1971). Written informed consent was obtained from all participants.

Analysis

The analysis procedure was the same as that used in Experiment 1, except that in Experiment 2 we did not carry out supplementary analyses on the response to the first noun phrase, since in this randomised design coordination and list trials were indistinguishable until the coordinator or placeholder was encountered. The exact rejection rates by condition for the critical epochs (time-locked to the coordinator) were: list-natural – 20%, list-jabberwocky – 24%, coordination-natural – 21%, coordination-jabberwocky – 21%.

Results

Behavioural

Overall accuracy on the memory task was 85.3% (natural list: 90.0%; natural coordination: 92.4%; jabberwocky list: 78.2%; jabberwocky coordination: 80.6%). A 2 × 2 repeated measure ANOVA revealed a main effect of content ($F(1,16) = 11.6$, $MSE = 2352.9$, $p < .01$), reflecting lower accuracy in the jabberwocky conditions. There was no main effect of structure ($F(1, 16) = 1.2$, $MSE = 94.1$, $p = .28$), and no interaction between structure and content ($F(1,28) = 0$, $MSE = 1$, $p > .2$).

ERPs

Visual inspection of the ERPs indicated that the results were qualitatively similar to Experiment 1, although slightly smaller in magnitude (Figures 3 and 4), and these observations were confirmed by statistical analysis (Table 2). In the 100–600 ms time-window (post-coordinator-onset), an omnibus quadrant ANOVA revealed a significant main effect of coordination ($F(1,16) = 15.4$, $MSE = 4.9$, $p < .01$), a significant interaction between coordination and hemisphere ($F(1,16) = 7.7$, $MSE = .64$,

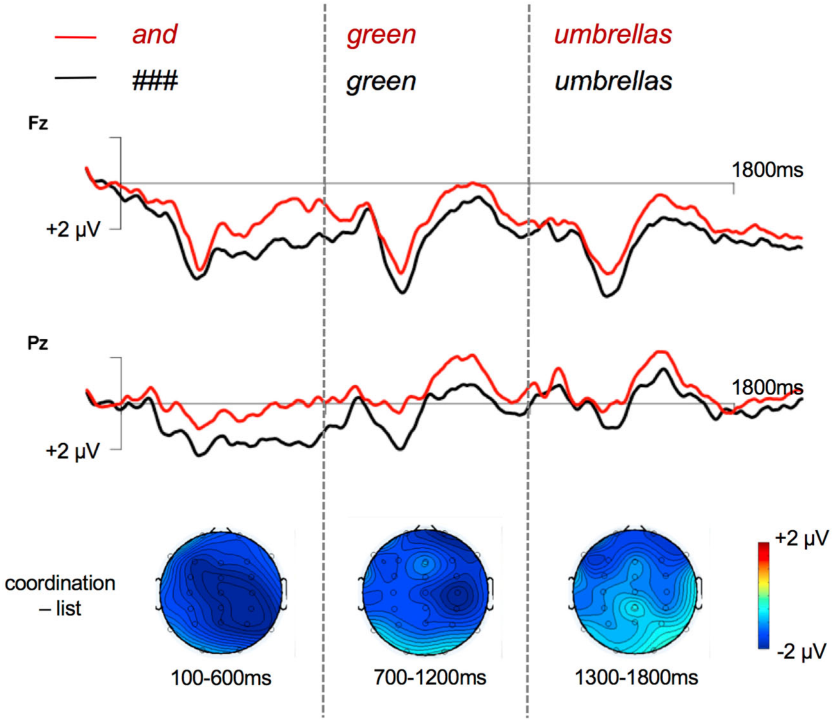

Figure 4. Scalp maps and event-related potentials time-locked to the coordinator (*and*) or the placeholder (*###*) at selected frontal and posterior electrode sites for natural list and natural coordination conditions in Experiment 2. Scalp maps illustrate the mean difference between coordination and list conditions across a 100–600 ms time-window following presentation of each word.

$p < .01$), and a significant 3-way interaction between coordination, hemisphere, and anteriority ($F(1,16) = 11.0$, MSE = .18, $p < .01$). The interaction between coordination and content also approached significance ($F(1,16) = 3.5$, MSE = 5.1, $p = .08$), reflecting a larger difference between coordinator and placeholder in the natural condition. The interaction between content, hemisphere, and anteriority was also marginally significant ($F(1,16) = 4.1$, MSE = .09, $p = .06$).

In the 700–1200 ms time-window at the adjective, we observed a marginal main effect of coordination ($F(1,16) = 4.5$, MSE = 7.5, $p = .05$), a significant interaction between coordination and content ($F(1,16) = 5.7$, MSE = 5.4, $p < .05$), and a significant interaction between coordination and hemisphere ($F(1,16) = 7.3$, MSE = .8, $p < .05$). As in Experiment 1, the interaction between coordination and content was driven by a large coordination effect in natural materials (1.4 µV) with little difference between jabberwocky conditions (.03 µV). In the subsequent 1300–1800 ms time-window at the noun, we again observed a significant interaction between coordination and content ($F(1,16) = 4.7$, MSE = 5.6, $p < .05$), again driven by a large coordination effect in natural materials (1.0 µV) and little difference between jabberwocky conditions (-.2 µV). We also observed a marginal interaction between coordination and anteriority ($F(1,16) = 3.2$, MSE = 1.2, $p = .09$).

Discussion

Experiment 2 was a replication of Experiment 1, this time randomising rather than blocking presentation of list and coordination trials. Despite the change in presentation, we observed essentially the same results as in Experiment 1: coordination resulted in a sustained anterior negativity across the course of the second noun phrase in natural conditions. Little difference between coordination and list trials was observed in jabberwocky conditions. These results provide further support for the effects of structural connectedness that we observed in Experiment 1.

It is interesting to note that in both experiments, the difference in the ERPs at the coordinator and the placeholder (###) position itself was numerically larger in the natural condition than in the jabberwocky condition. Some speculative explanations for this trend which could be pursued in future work are (a) early part of the structure maintenance cost; (b) simple attentional differences between jabberwocky and natural stimuli; (c) additional predictive processing for the content or structure of the upcoming noun phrase in the coordination case.

General discussion

We conducted two ERP experiments designed to investigate the hypothesis that sustained neural activity plays a role in encoding linguistic structure across the course of a phrase or sentence. We looked for evidence of increased activity while processing a second noun phrase that was syntactically coordinated with a previous noun phrase, relative to a case in which the second noun phrase was just the second element in an uncoordinated list. In both experiments we found a sustained increased anterior negativity across the second noun phrase for coordinated trials when natural materials were used, but not when jabberwocky materials were used.

Active maintenance of syntactic structure

As reviewed in the introduction, Pallier et al. (2011), following from Smolensky and Legendre's (2006) computational model of syntactic structure encoding, proposed that each new terminal added to a current connected syntactic representation increases the amount of neural activity required to maintain the entire representation. Pallier et al.'s (2011) account predicted that neural activity associated with syntactic structure encoding should thus increase monotonically across the course of the largest syntactic constituent, but the poor temporal resolution of fMRI data made it difficult to precisely evaluate the timing of the increased activity that they observed. The results of the current ERP study showed a sustained difference between structured and unstructured conditions with the temporal profile predicted by this hypothesis when using natural materials with lexical semantic content. In the list condition, there was no syntactic connection between the first noun phrase and the second noun phrase, and therefore after the first noun phrase the syntactic "buffer" would be cleared and structural encoding would restart on the second noun phrase. In the coordination condition, the coordinator indicated that both noun phrases belong to a single connected syntactic representation. Therefore, as the second noun phrase was processed, the size of the syntactic representation being maintained increased, which would result in increased neural activity during this period relative to the list condition. We note that we did not observe any increase in the size of the negativity across the course of the second noun phrase, which might seem to be predicted by the Pallier et al. (2011) hypothesis as e.g. seeing the adjective will increase the number of nodes being maintained in memory; however, as their proposed neural activation function was logarithmic, a longer sequence might be required to adequately evaluate this prediction.

The difference between conditions during the second noun phrase took the form of a negativity with a focus at anterior electrodes. This is somewhat consistent with the broader ERP literature, where it has long been suggested that sustained anterior negativities reflect working memory processes in sentence comprehension (Fiebach et al., 2002; Kluender & Kutas, 1993a, 1993b; Münte, Schiltz, & Kutas, 1998; Phillips et al., 2005; Ueno & Kluender, 2003). However, these prior studies focused on different operations than simply encoding prior structure; many examined the narrower hypothesis that syntactic predictions in particular require active maintenance, such as the "filler" in wh-filler gap constructions, while Münte et al. suggested that the interpretation of temporally reversed clauses (*before x, y*) required extra working memory processing – a semantic operation related to ongoing integration rather than a syntactic operation related to maintenance (Münte et al., 1998). It is also not clear how these differences in overall evoked power relate to the differences between sentences and word lists observed for specific frequency bands such as theta, beta, or high gamma (Bastiaansen et al., 2010; Fedorenko et al., 2016). Therefore much more work is needed to determine whether these effects in fact reflect a common process, and to develop a theory of what exactly that process could be. Some promising initial steps in this direction are evident in new results from Nelson et al. (2017), who obtained direct intracranial recordings from patients reading sentences of various lengths and structures; they report that high gamma power was sensitive to the number of syntactic nodes that still belong to an incomplete constituent (maintenance of prior structure, modulo compression of finished sub-phrases).

Pallier et al. (2011) and Nelson et al. (2017) report constituency-size effects for syntax in left inferior frontal and left posterior superior temporal cortices. Indeed, a very long and storied literature has associated left inferior frontal cortex with working memory and integration processes, with much debate over whether these processes are language- or syntax-specific (Caplan & Waters, 1999; Fedorenko et al., 2010; Fiebach, Schlesewsky, Lohmann, Von Cramon, & Friederici, 2005; Hagoort, 2005). Other work has suggested that posterior temporal cortex might serve in short-term memory mechanisms akin to the classic phonological loop, in which phonological or lexical representations in posterior temporal cortex are cyclically reactivated in order to maintain them in short-term memory (Buchsbaum & D'Esposito, 2008; Hickok, 2009; Hickok & Poeppel, 2007). Some fMRI studies that used long-distance dependency paradigms of the kind that elicited sustained anterior negativities in EEG reported increased activity in inferior frontal cortex (e.g. Fiebach et al., 2005; Santi & Grodzinsky, 2007) but others have not (Matchin, Sprouse, & Hickok, 2014), and several recent studies have associated this activity with resolution of interference rather than active maintenance of syntactic structure (Glaser et al., 2013; Santi, Friederici, Makuuchi, & Grodzinsky, 2015). Although the poor spatial resolution of EEG would make it difficult to precisely determine the source of the anterior negativity in the current study, future work with MEG will provide better estimates.

A different question raised by evidence for active maintenance of linguistic structure is whether these mechanisms result in better memory for the sub-elements of the structure. In the current study, we did not observe significantly better memory recognition performance for phrases when they were coordinated rather than separated by a placeholder, although there was a trend in this direction in Experiment 2 when a randomised presentation was used. However, this study was not designed to test behavioural effects of structure and therefore we only collected memory recognition responses on a small number (10%) of trials. Examining more systematically the possible effects of active structure maintenance on memory for sub-parts of the structure will be an interesting direction for future work.

Dissociating syntactic and semantic effects

Although the sustained anterior negativity observed in the natural conditions is consistent with Pallier et al.'s syntactic encoding hypothesis, the failure to observe these effects in the jabberwocky conditions raises the alternative possibility that the anterior negativity resulted from differences in semantic processing. Dissociating syntactic from semantic structure-building neural processes is well-known to be challenging because the complexity of the compositional semantic representation is obviously highly correlated with the complexity of the syntactic representation (see Pylkkänen, Brennan, & Bemis, 2011, for more discussion). Pallier et al. (2011) used constituency-size effects in jabberwocky materials in an attempt to dissociate syntactic and semantic encoding, arguing that regions demonstrating constituency-size effects for materials lacking lexical semantics must be involved in syntactic encoding. If the sustained anterior negativity observed in the natural conditions in the current study reflects semantic rather than syntactic processes, these data would not provide any positive support for the Pallier et al. active maintenance syntactic encoding account, but they would suggest that small functional elements can have a surprisingly large and automatic impact on semantic processing even outside of full sentence contexts.

In order to maintain the syntactic maintenance account of the anterior negativity in the natural conditions, we would need to assume that participants simply failed to assign syntactic structure to the jabberwocky items in the current experiment. One reasonable explanation for this would be the limited functional cues to structure in the current paradigm. Here the only functional cues in the jabberwocky condition were the use of frequent adjectival endings (*sunc**y***) to cue that the first word was an adjective, the use of the plural –s to cue that the second word was a noun (*kad**s***), and the coordinator itself. However, these limited cues may not have been as effective at inducing parsing of jabberwocky as the more diverse cases used by Pallier et al. (2011) in which less ambiguous functional cues may have appeared earlier in the string (e.g. *in the suncy kads*). Future work could explore this possibility by testing materials with additional functional material (e.g. *the suncy kads and the whisy cawregrols*).

Semantic accounts of the structural connectedness effects observed here could fall into two categories. On the one hand, one could assume an active structure maintenance account functionally parallel to the Pallier et al. (2011) proposal, but specified for semantic rather than syntactic representations. In other words, the sustained anterior negativity would reflect the binding of lexical-semantic information to some structured semantic object (e.g. logical form, or a propositional representation) and the maintenance of this binding as this object is constructed across the course of the phrase. Pallier et al. suggested this kind of account for their observation that anterior temporal cortex regions showed constituency-size effects in natural but not jabberwocky materials. Similarly, in ECoG patients Fedorenko et al. (2016) found a monotonic increase in gamma power across the course of regular sentences but not for jabberwocky sentences, and attributed this increase to the representation of complex meanings.

Alternatively, the sustained anterior negativity might rather be attributed to other semantic processes that might differ as a function of syntactic context, such as conceptual integration and discourse representation. For example, Pallier et al. (2011) speculated that linguistic input might only be integrated into the discourse when it forms complete sentences, and suggested that this might explain why a temporoparietal region showed increases only for conditions with longer phrases. The current study did not use complete sentences, but several colleagues have noted their subjective impressions that the presence of the coordinator makes one more likely to attempt to imagine a scenario in which both of the referents described would be discussed together, to imagine a predicate which might apply to both, or to visualise both. We are not aware of previous work associating anterior negativities with

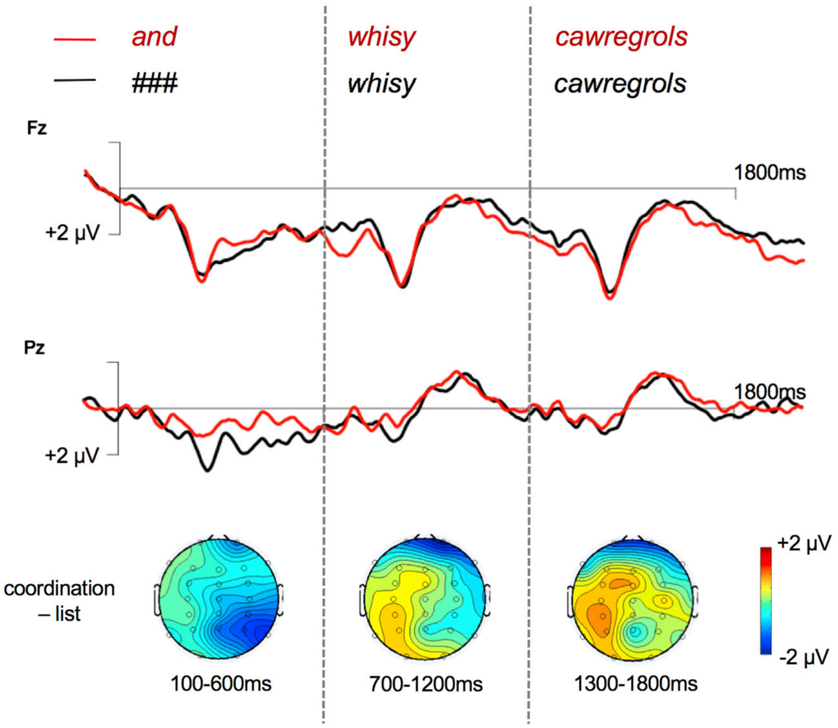

Figure 5. Scalp maps and event-related potentials time-locked to the coordinator (*and*) or the placeholder (*###*) at selected frontal and posterior electrode sites for jabberwocky list and jabberwocky coordination conditions in Experiment 2. Scalp maps illustrate the mean difference between coordination and list conditions across a 100–600 ms time-window following presentation of each word.

these kinds of processes in particular, but several studies suggest that anterior negativities are associated with "frame-shifting" or reinterpretation as required for processing figurative language (e.g. Coulson & Kutas, 2001; Wlotko & Federmeier, 2012). Future studies might directly test the latter kind of alternative account by systematically varying the imageability and conceptual association between the two noun phrases and evaluating whether there is any impact on the sustained anterior negativity for structure.

Conclusion

We reported two ERP experiments that investigated neural mechanisms for encoding linguistic structure by measuring the impact of coordination on the neural response to the coordinated elements. We found that processing a noun phrase as part of a larger connected coordinate structure is associated with a differential neural response – a sustained anterior negativity – throughout the noun phrase, relative to the unstructured comparison case. Although much work remains to be done to distinguish alternative accounts of the underlying mechanisms, these results crucially show that structural connectedness exerts strong, ongoing effects on neural activity even when the properties of structured and unstructured materials are very tightly controlled.

Acknowledgments

We would like to thank William Matchin, Chris Hammerly, Natalia Lapinskaya, Mina Hirzel, and especially Anna Namyst for assistance with data collection and analysis.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Bastiaansen, M., Magyari, L., & Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *Journal of Cognitive Neuroscience*, *22*(7), 1333–1347.

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

Buchsbaum, B. R., & D'Esposito, M. (2008). The search for the phonological store: From loop to convolution. *Journal of Cognitive Neuroscience*, *20*(5), 762–778.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, *22* (01), 77–94.

Coulson, S., & Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters*, *316*(2), 71–74.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21.

Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, *19*(1), 158–164.

Fedorenko, E., Hsieh, P.-J., Nieto-Castañon, A., Whitfield-Gabrieli, S., & Kanwisher, N. (2010). New method for fMRI investigations of language: Defining ROIs functionally in individual subjects. *Journal of Neurophysiology*, *104*, 1177–1194.

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

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language*, *47*(2), 250–272.

Fiebach, C. J., Schlesewsky, M., Lohmann, G., Von Cramon, D. Y., & Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: Syntactic integration versus syntactic working memory. *Human Brain Mapping*, *24*(2), 79–91.

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.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology*, *61*(2), 331–349.

Fuster, J. M., & Alexander, G. E. (1971). Neuron activity related to short-term memory. *Science*, *173*(3997), 652–654.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*(1), 1–76.

Glaser, Y. G., Martin, R. C., Van Dyke, J. A., Hamilton, A. C., & Tan, Y. (2013). Neural basis of semantic and syntactic interference in sentence comprehension. *Brain and Language*, *126*(3), 314–326.

Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*(9), 416–423.

Hickok, G. (2009). The functional neuroanatomy of language. *Physics of Life Reviews*, *6*(3), 121–143.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393–402.

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.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627–633.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*(1), 15–47.

King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience*, *7*(3), 376–395.

Kluender, R., & Kutas, M. (1993a). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience*, *5*(2), 196–214.

Kluender, R., & Kutas, M. (1993b). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*(4), 573–633.

Lewis, R., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, *29*(3), 375–419.

Lewis-Peacock, J. A., Drysdale, A. T., Oberauer, K., & Postle, B. R. (2012). Neural evidence for a distinction between short-term memory and the focus of attention. *Journal of Cognitive Neuroscience*, *24*(1), 61–79.

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*.

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., & Miller, E. K. (2016). Gamma and beta bursts underlie working memory. *Neuron*, *90*(1), 152–164.

Matchin, W., Hammerly, C., & Lau, E. (2017). The role of the IFG and pSTS in syntactic prediction: Evidence from a parametric study of hierarchical structure in fMRI. *Cortex; a journal devoted to the study of the nervous system and behavior*, *88*, 106–123.

Matchin, W., Sprouse, J., & Hickok, G. (2014). A structural distance effect for backward anaphora in Broca's area: An fMRI study. *Brain and Language*, *138*, 1–11.

Mazoyer, B. M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., … Mehler, J. (1993). The cortical representation of speech. *Journal of Cognitive Neuroscience*, *5*(4), 467–479.

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*(1), 67–91.

Miller, G. A., & Isard, S. (1964). Free recall of self-embedded English sentences. *Information and Control*, *7*(3), 292–303.

Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature*, *395*(6697), 71–73.

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

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

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

Phillips, C., Kazanina, N., & Abada, S. (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, *22*(3), 407–428.

Postle, B. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience*, *139*(1), 23–38.

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

R Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrived from http://www.R-project.org/

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

Ruchkin, D. S., Johnson, R., Canoune, H., & Ritter, W. (1990). Short-term memory storage and retention: An event-related brain potential study. *Electroencephalography and Clinical Neurophysiology*, *76*(5), 419–439.

Ruchkin, D. S., Johnson, R., Grafman, J., Canoune, H., & Ritter, W. (1992). Distinctions and similarities among working memory processes: An event-related potential study. *Cognitive Brain Research*, *1*(1), 53–66.

Sachs, J. S. (1967). Recopition memory for syntactic and semantic aspects of connected discourse. *Attention, Perception, & Psychophysics*, *2*(9), 437–442.

Santi, A., Friederici, A. D., Makuuchi, M., & Grodzinsky, Y. (2015). An fMRI study dissociating distance measures computed by Broca's area in movement processing: Clause boundary vs. identity. *Frontiers in Psychology*, *6*.

Santi, A., & Grodzinsky, Y. (2007). Working memory and syntax interact in Broca's area. *NeuroImage*, *37*(1), 8–17.

Smolensky, P., & Legendre, G. (2006). *The harmonic mind: From neural computation to optimality-theoretic grammar*. Cambridge, MA: MIT Press.

Stokes, M. G. (2015). 'Activity-silent' working memory in prefrontal cortex: A dynamic coding framework. *Trends in Cognitive Sciences*, *19*(7), 394–405.

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

Ueno, M., & Kluender, R. (2003). Event-related brain indices of Japanese scrambling. *Brain and Language*, *86*(2), 243–271.

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.

Van Dyke, J., & Johns, C. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass*, *6*(4), 193–211.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class words. *Memory and Cognition*, *19*(1), 95–112.

Wlotko, E. W., & Federmeier, K. D. (2012). So that's what you meant! Event-related potentials reveal multiple aspects of context use during construction of message-level meaning. *NeuroImage*, *62*(1), 356–366.

Yngve, V. H. (1960). A model and an hypothesis for language structure. *Proceedings of the American Philosophical Society*, *104*(5), 444–466.