Syntactic Complexity and Frequency in the Neurocognitive Language System

Yun-Hsuan Yang[1,2], William D. Marslen-Wilson[1,2],
and Mirjana Bozic[1,2]

Abstract

■ Prominent neurobiological models of language follow the widely accepted assumption that language comprehension requires two principal mechanisms: a lexicon storing the sound-to-meaning mapping of words, primarily involving bilateral temporal regions, and a combinatorial processor for syntactically structured items, such as phrases and sentences, localized in a left-lateralized network linking left inferior frontal gyrus (LIFG) and posterior temporal areas. However, recent research showing that the processing of simple phrasal sequences may engage only bilateral temporal areas, together with the claims of distributional approaches to grammar, raise the question of whether frequent phrases are stored alongside individual words in temporal areas. In this fMRI study, we varied the frequency of words and of short and long phrases in English. If frequent phrases are indeed stored, then only less frequent items should generate selective left frontotemporal activation, because memory traces for such items would be weaker or not available in temporal cortex. Complementary univariate and multivariate analyses revealed that, overall, simple words (verbs) and long phrases engaged LIFG and temporal areas, whereas short phrases engaged bilateral temporal areas, suggesting that syntactic complexity is a key factor for LIFG activation. Although we found a robust frequency effect for words in temporal areas, no frequency effects were found for the two phrasal conditions. These findings support the conclusion that long and short phrases are analyzed, respectively, in the left frontal network and in a bilateral temporal network but are not retrieved from memory in the same way as simple words during spoken language comprehension. ■

INTRODUCTION

The dominant view of language processing holds that linguistic knowledge is built around two separate mechanisms (e.g., Pinker, 1999; Chomsky, 1965): a syntactic combinatorial mechanism that assembles syntactic hierarchical structures from lexical items stored in a second mechanism, the lexicon. The assembling of phrases or sentences requires syntactic operations, where sequences of words (or morphemes) are combined together to build hierarchical structures. Given that all the words in a given sequence—for example, "go to a concert"—exist in the lexicon and that the meaning of this sequence can be unambiguously computed by syntactic combinatorial operations, there will be no separate representation in the lexicon of "go to the concert" (or of any comparable multiword sequence). Syntactic parsing is seen as an obligatory process, so that regardless of how many times a phrase has been heard or produced, its interpretation will still require syntactic analysis.

In line with this view, prominent proposals for the neural mechanisms underlying language comprehension generally posit two mechanisms, a morphosyntactic parsing mechanism and a mechanism for access to lexical form and content. These theories suggest that a left hemisphere (LH) network involving dorsal parts of the LIFG (BA 44/45) and posterior superior and middle temporal gyri make up a core left frontotemporal mechanism for complex syntactic structure processing, functionally integrated by two major white matter tracts, the arcuate fasciculus and the extreme capsule (Griffiths, Marslen-Wilson, Stamatakis, & Tyler, 2013; Rolheiser, Stamatakis, & Tyler, 2011). In contrast, an overlapping but more ventral network, involving the left superior temporal gyrus (STG), middle temporal gyrus (MTG), and inferior regions of the LIFG (BA 47), supports the access of meanings of words (Hagoort & Indefrey, 2014; Hagoort, 2005, 2013; Friederici, 2011). Other proposals contend that the ventral network is more bilaterally organized and includes both left and right inferior frontal gyri (IFG) and temporal areas (Bozic, Tyler, Ives, Randall, & Marslen-Wilson, 2010; Hickok, 2009; Tyler & Marslen-Wilson, 2008). Despite these differences, it is usually agreed that the BA 44/45 in left inferior frontal cortex and linked posterior left temporal areas are essential for hierarchical syntactic processing. This strict separation between a syntax parser and a lexicon presupposes that the parsing of all syntactically complex items is obligatory.

Consistent with the predictions of this type of account, several neuroimaging studies have shown that morphosyntactically complex words, phrases, and sentences engage the left frontotemporal network (e.g., Bozic et al., 2010; Friederici, Fiebach, Schlesewsky, Bornkessel, & von

[1]University of Cambridge, [2]MRC Cognition and Brain Sciences Unit, Cambridge, UK

Cramon, 2006). Lesion studies also confirm that the LIFG and posterior left temporal areas are indispensable for morphosyntactic processing (e.g., Wright, Stamatakis, & Tyler, 2012; Tyler et al., 2011). Many of the studies, however, that report a greater activation of the LIFG during syntactic processing tend to use highly syntactically complex sentences involving center-embedding (Makuuchi, Bahlmann, Anwander, & Friederici, 2009) or scrambling (e.g., Friederici et al., 2006). On the other hand, results from studies using syntactically simple and canonical utterances seem to suggest that such strings need only involve bilateral temporal structures (Tyler et al., 2010; Friederici, Wang, Herrmann, Maess, & Oertel, 2000), consistent with recent evidence that minimally syntactically complex items such as short phrases (e.g., "I play") do not selectively engage the left frontotemporal network but instead are associated with bilateral temporal activation (Bozic, Fonteneau, Su, & Marslen-Wilson, 2015).

The Bozic et al. (2015) finding in particular poses a challenge to the standardly assumed clear demarcation between lexicon and syntax and may be more consistent with the predictions made by distributional approaches to grammar (e.g., Bybee & McClelland, 2005). Such approaches do not presuppose a strict division between lexical items and syntactic constructions in the same way as generative grammar. On this view, linguistic items ranging from morphemes to words to syntactically more complex structures are all potentially storable as exemplars, which can be retrieved during language processing. Although this is not the interpretation offered by Bozic et al. (2015), this perspective could account for the bilateral activation for simple phrases observed in their data. It is plausible that the two-word sequences investigated in this study were frequently encountered phrases ("I play" or "the dog"), which may be neurally stored as complete memory traces. This would allow listeners to directly retrieve their meaning in the same way as simple words, via the bilateral network.

Frequency Effects during Language Processing

It is well documented that language users are sensitive to the frequency of individual words (e.g., Dahan, Magnuson, & Tanenhaus, 2001). More relevantly, effects of frequency are also present for linguistic units larger than single words, raising the possibility that the use of certain syntactic constructions is not necessarily evidence that a highly abstract syntactic pattern is at work. For instance, recent behavioral studies show that language users do have knowledge about the frequency of specific phrases: Arnon and Snider (2010) showed that more frequent multiword sequences (such as "don't have to worry") are recognized faster than less frequent ones (such as "don't have any place"). Janssen and Barber (2012) found that the naming latency of two- and three-word phrases decreases as a function of the frequency of the phrase as a whole. Overall, there is solid behavioral evidence suggesting an influence of frequency information above the level of individual words.

Although distributional approaches to grammar can account for the frequency effects found for phrases, as well as the Bozic et al. (2015) findings, they do not easily square with the neuroimaging results showing a functional partition between the processing of syntactically complex sentences in the left frontotemporal network and simple words in more temporally distributed regions (Hagoort, 2014; Friederici, 2011). Given the research discussed above, it is possible that both approaches are partially right. In line with distributional approaches to grammar, the bilateral network may store heard words, phrases, and sentences as a function of their frequency of occurrence. However, in addition to this system, we still need a separate mechanism for morphosyntactic processing (or "unification"; cf. Hagoort, 2014) of complex or infrequent structures—namely the left frontotemporal network.

During spoken language comprehension, therefore, the functional partition between the two language networks may be codetermined by the syntactic complexity and the frequency of occurrence of a specific input. When the encountered item is highly frequent, such as "I think," the listener will have a representation in the bilateral network readily available so that they can simply retrieve its meaning from memory. On the other hand, when the item is lower in frequency, the left frontotemporal network will be recruited because a memory trace for the phrase has not been established. On this basis, the bilateral temporal mechanism for processing simple words will also support the mapping between sound and meaning for frequently encountered phrases and sentences. If so, then these regions should show modulation by frequency for words and phrases, whereas the left frontotemporal network will primarily be engaged by lower frequency complex structures.

The existing literature offers substantial evidence that frequency of single words affects the activation of language-related brain areas. For instance, in the visual domain, lower frequency words activated LIFG (BA 44/45), fusiform gyrus, and anterior cingulate regions more strongly than high-frequency words (Chee, Westphal, Goh, Graham, & Song, 2003). In an fMRI study using lexical decision, Nakic, Smith, Busis, Vythilingam, and Blair (2006) found that low-frequency words generated more activity in bilateral IFG regions (left BA 45 and right BA 47) compared with high-frequency words. However, to our knowledge, no neuroimaging studies have tested for sensitivity to the frequency of phrases and sentences. This study therefore looks at the potential effects of frequency across different levels of syntactic complexity on the activation of the cortical language networks.

The Present Study

The goal of this experiment was to investigate the potential interplay between the effects of frequency and syntactic

complexity on patterns of neural activation during the interpretation of spoken words and short sequences of words. In doing so, we will evaluate three hypotheses. First, on a strong storage-based account, the frequency of syntactically complex forms (in this experiment two- and three-word phrases) will correlate with degree of activation of the core LH network for combinatorial syntax. The more frequent the phrases, the more likely that they will have a stored representation available in middle temporal regions, predicting less engagement of this core frontotemporal combinatorial substrate, especially where dorsal LIFG (BA 44/45) is concerned. On a strong obligatory syntactic parsing account, in contrast, frequency of occurrence will not modulate the activation of this network, and both two- and three-word sequences will be equally effective in triggering selective LIFG activity. Finally, the Bozic et al. (2015) approach, though not taking a stand on the role of sequence frequency, does suggest a difference between minimal two-word sequences and more syntactically complex three-word sequences. Only the latter should selectively engage the core LH system.

We examined these hypotheses in an event-related fMRI experiment which covaried the two primary variables of syntactic complexity and frequency of occurrence. For syntactic complexity, taking single words (verbs and nouns without inflectional morphemes) as a baseline (Figure 1A, D), we added two further levels of complexity: "short phrases" and "long phrases." The short phrases (Figure 1B, E) are simple two-word phrases

with a flat structure, with a dependency relationship between two-word categories (e.g., "I act," "our shop"). These are comparable to the minimally complex two-word phrases tested by Bozic et al. (2015). Long phrases (Figure 1C, F) are constructed by building an extra structure onto a three-word phrase, resulting in a hierarchical structure (e.g., "I support it," "on a tree"). None of the phrases included inflectional affixes.

For each level of syntactic complexity, we selected items varying in corpus frequency from low through mid to high, with frequency of low/mid/high items closely matched across the three complexity levels. Note that the frequency manipulation here refers to the frequency of occurrence of the whole word for simple words and to the frequency of the whole phrase for short and long phrases. To ensure that any differential frequency effects for the short and long phrases can be attributed to this manipulation of the whole-phrase frequency, rather than the frequency of its constituent words, we also matched the frequency of the main verbs and nouns in short and long phrases.

The paradigm was designed to make the listening environment in the scanner as natural as possible and to minimize potentially misleading LIFG activations generated by task requirements (cf. Wright, Randall, Marslen-Wilson, & Tyler, 2011). Drawing on the method employed by Bozic et al. (2015), a seminatural listening paradigm was used in which participants were simply asked to listen attentively to auditory stimuli and to perform an occasional semantic judgment task (5% of the trials) to keep them alert. Data were analyzed using a combination of univariate and multivariate methods to assess both the overall amplitude differences between conditions and the information coded in the relationship between activated voxels across conditions. More specifically, we used standard univariate analyses to reveal the extent of activation across conditions in the language processing network, and multivariate Representational Similarity Analysis (RSA, Kriegeskorte, Goebel, & Bandettini, 2006; Nili et al., 2014) to explore correlations between these condition-based activation patterns and different theoretical models, in order to reveal the specific processes that they encode. This combination of approaches is arguably necessary to disentangle a multitude of concurrent subprocesses that contribute to the activation in frontotemporal language-related brain regions (e.g., Hagoort & Indefrey, 2014). Activations for each level of syntactic complexity were examined by first subtracting out the matched acoustic baseline MuR (Musical Rain; cf. Uppenkamp, Johnsrude, Norris, Marslen-Wilson, & Patterson, 2006). The effect of frequency on each level of syntactic complexity was investigated using parametric modulator analyses to look for brain regions in which activation linearly increases or decreases with frequency, followed by item-based RSA tests of the spatial patterns associated with frequency at different levels of syntactic complexity.
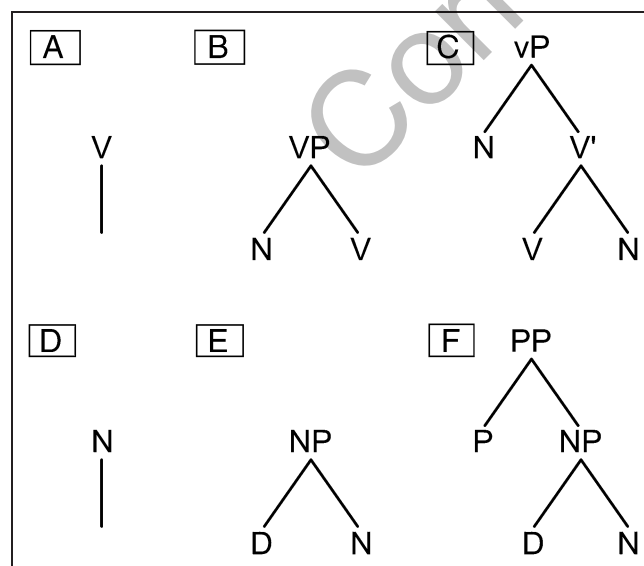

Figure 1. Syntactic trees for three levels of syntactic complexity. Trees plotted separately for the verbal (upper series) and nominal (lower series) conditions: (A) simple verb, (B) short verb phrase, (C) long verb phrase, (D) simple noun, (E) short noun phrase, and (F) long noun phrase.

METHODS

Participants

Eighteen right-handed native speakers (10 men) of British English with normal hearing were recruited from the University of Cambridge community. The participants were classified as monolingual on the basis that they reported that they did not speak another language to a native or near-native level. The age range of the participants was 19–33 years (mean age = 24.7 years). They had no known hearing, language, or neurological impairments and had normal or corrected-to-normal vision. The study was approved by Cambridge Psychology research ethics committee.

Materials and Design

Two main variables, syntactic complexity and frequency, were contrasted in a 3 × 3 design. The three levels of syntactic complexity were word, short phrase, and long phrase. The three levels of frequency were low, mid, and high. Sixty trials were included in each condition (Table 1).

The stimuli for each condition consisted of 50% verbal and 50% nominal structural types. The word condition included simple verbs and nouns (e.g., "implore," "sequin"); the verb short phrase and noun short phrase conditions were "subject pronoun (I, you, we) + verb" (e.g., "I visit") and "possessive pronoun (My, your, our) + noun" (e.g., "your plan"); the verb long phrase and noun long phrase conditions were "subject pronoun (I, you, we) + verb + object pronoun (it, you, me)" (e.g., 'I love it') and "preposition (as, at, for, in, on, to) + possessive pronoun or determiner (my, your, our, the, a) + noun" (e.g., "on the phone"; Figure 1). Fifteen native speakers of English who did not take part in the main experiment rated the naturalness of the stimuli on a scale from 1 to 7 and short and long phrases in the experiment received an average rating above 3.5. All the verbs and nouns used are verb dominant and noun dominant items, respectively.

Low-frequency items had a frequency of 0.03–0.3 per million words, mid frequency items had 0.31–0.79, and high-frequency items had 0.8–10 per million words, as retrieved from the British National Corpus (Table 1). Frequency was controlled for across the three complexity levels: There were no significant frequency differences between the simple word, short phrase, and long phrase conditions within the same frequency band ($p > .05$). In addition, the frequency of the main verb and noun in short and long phrase conditions was also controlled ($p > .05$). This was to make sure that any differences found between phrases like "I order" and "I support it" could not be attributed to variations in lexical retrieval difficulty associated with the main verb or noun.

It is important to note, given the requirement to match frequency across syntactic complexity conditions, that the frequency of the simple words was constrained by the frequency of short and long phrases. These do not occur with anything like the level of frequency found for single lexical items, with the result that the single word items have frequency levels that are low in comparison with the levels typically used in the word frequency literature. This was unavoidable if frequency was to be matched across words, short phrases and long phrases within each of the three levels of frequency manipulation. For the short phrase and long phrase conditions, the most frequent items available in the corpus were selected.

The stimuli were recorded by a female native speaker of Southern British English, digitized at 44.1 kHz, and then downsampled to 22 kHz. There were 60 lexical trials in each condition, for a total of 540 trials. In addition, 20 items were randomly selected from each condition to provide the templates for the MuR acoustic baseline trials (Uppenkamp et al., 2006). These preserve the temporal duration, the temporal envelope, and the energy levels of the original speech stimuli. This resulted in 60 MuR trials for each level of syntactic complexity, for a total of 180 MuR trials. The design also included 160 silence trials and 40 one-back semantic task trials. All trials were evenly split into five blocks, each lasting about 15 min, and pseudorandomized, such that the same condition did not appear more than twice in a row.

Procedure

A seminatural listening paradigm was used, in which participants heard stimuli via headphones and were asked to perform an infrequent (5% of trials) 1-back semantic judgment task. The participants had to decide whether a word (adjective or adverb), which appeared on the screen, could be meaningfully related to the previously heard stimulus (e.g., "impersonate"–"amazingly" and "we find it"–"easily"). Each trial started with 100 msec of fixation cross with no auditory presentation, after which a

Table 1. Conditions in the Experiment and Example Stimuli for Both the Verbal and Nominal Items

Syntactic Complexity	Frequency		
	Low (0.03–0.3 per Million; M = 0.17)	Mid (0.31–0.79 per Million; M = 0.51)	High (0.8–10 per Million; M = 4)
Low (simple word)	inscribe; sequin	oppress; turnip	pretend; spinach
Mid (short phrase)	I order; our shop	I answer; my report	I guess; my plan
High (long phrase)	I support it; for my child	I remember you; in my garden	I love it; on the phone

stimulus was delivered through the headphones. On the task trials, the 100-msec fixation was followed by a word in the middle of the screen, to which participants had to respond by pressing buttons using the index or middle finger of their right hand. There was a short break after the third block for the participants when T2 and fieldmaps were acquired. Before scanning, participants did a short practice session outside the scanner room. They also underwent a short sound test in the scanner to make sure the sound input was balanced between the two ears and the headphones were installed correctly. The sounds were delivered with NNL Electrostatic headphones using E-Prime presentation software. The total duration of the experiment including breaks and structural scans was around 1 hr 15 min.

Data Acquisition

The data were acquired with a 3-T Trio Siemens (Erlangen, Germany) scanner at MRC Cognition and Brain Sciences Unit, Cambridge. Fast-sparse gradient-echo EPI sequence was used to remove the effect of EPI noise during the presentation of auditory stimuli (repetition time = 3.4 sec, acquisition time = 2 sec, echo time = 30 msec, flip angle = 78°, matrix size = 64 × 64, field of view = 192 × 192 mm, 32 oblique slices 3 mm thick, 0.75 mm gap). T1-weighted structural scans were obtained for anatomical localization (3-D MPRAGE sequence; repetition time = 2250 msec, echo time = 2.99 msec, flip angle = 9°, field of view = 256 × 240 × 192 mm, matrix size = 256 × 240 × 192 mm, spatial resolution 1 mm isotropic).

Data Analyses

Preprocessing was done using Automated Analysis version 4 and SPM 8 (Cusack et al., 2014). The preprocessing

steps included image realignment (for movement correction), segmentation, spatial normalization to the MNI standard brain, and smoothing using a 10-mm Gaussian kernel. No slice timing correction was employed because the sparse-sampling imaging acquisition used in this study could render interpolation inaccurate.

Univariate Analyses

The univariate statistical analyses were carried out using the general linear model with five blocks and 38 event types (18 verbal test conditions, 18 MuR conditions, 1 task, 1 silence). The neural response was modeled with the canonical hemodynamic response function. Six motion regressors were included to separate out the effects of movement. A high-pass filter with a 128-sec cutoff was used to remove low-frequency noise. Group data were analyzed using random effects analsyis. The results were thesholded at a voxel level of $p < .001$ and a cluster level of $p < .05$ (corrected for multiple comparisons).

Multivariate Analyses

To run the multivariate analyses, we used the searchlight RSA procedure as implemented in the RSA toolbox (Nili et al., 2014). Central to searchlight RSA are comparisons between observed similarity in spatial patterns of neural activity between conditions, represented as data representational dissimilarity matrices (RDMs), and model RDMs, which are constructed to reflect competing hypotheses about how the conditions differ in their neural patterns (see schematic diagram of RSA procedure in Figure 2). Each RDM is a matrix where each cell is a similarity value, namely a correlation coefficient between a pair of conditions, observed (as in data RDMs) or

Figure 2. Procedure for searchlight RSA analysis. A 5-mm radius searchlight (A) moves voxel by voxel across the brain to collect voxel activation patterns associated with each condition (B). The pairwise correlational similarity between the voxel patterns for each condition is used to construct the 18 × 18 data RDM, computed for each searchlight stopping point (C). Each such data RDM is then correlated with each model RDM to build a brain-wide map of model fit (D). Panel (E) illustrates the correlational structure assumed by the example model RDM—highly correlated patterns within a condition (blue cells) but only weak correlations between conditions (red cells).
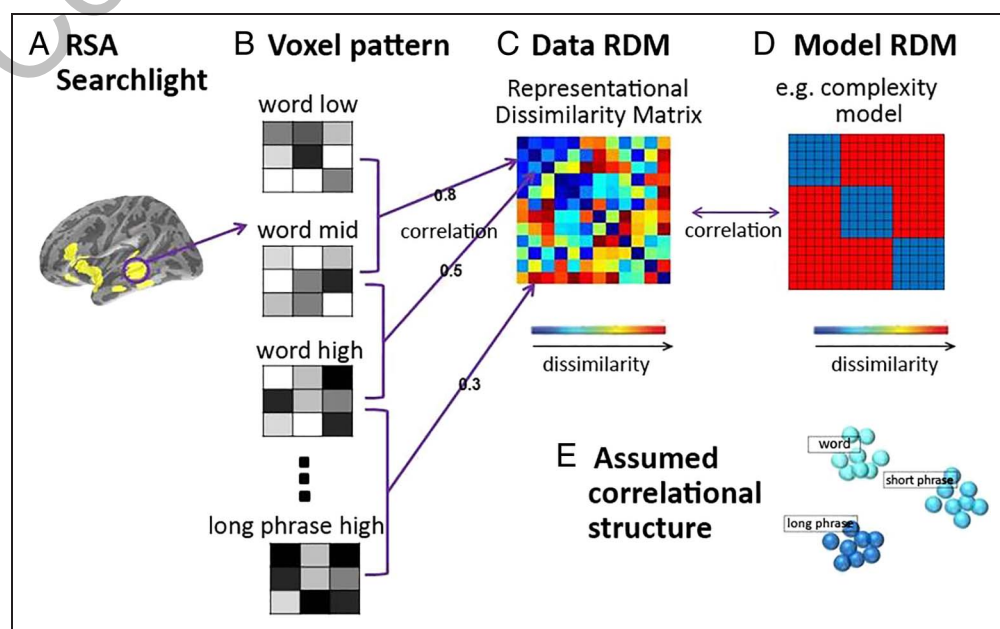
hypothesized (as in model RDMs). The degree of similarity between data RDMs and model RDMs is revealed by calculating a correlation coefficient, and the value is associated with the voxel at the center of the searchlight.

For RSA, the conditions were first modeled (using general linear model) as epochs, corresponding to the duration of the sound file of each item. The resulting beta values for the conditions at each voxel were then submitted to searchlight RSA, where a 5-mm radius spherical searchlight was moved voxel by voxel across the whole brain of each participant, collecting the beta values for the voxels falling within each searchlight stopping point (Figure 2). Voxel patterns at each searchlight location were then correlated for each pair of conditions, yielding data RDMs across all voxels. The data RDMs were then compared with model RDMs, which we constructed to represent predictions based on our theoretical hypotheses about how the conditions should cluster together in the underlying neural space (see below). This resulted in a correlation map for each participant to reveal regions where activation patterns correspond to the model RDMs. Finally, group level statistical tests were performed at each voxel by a signed-rank test across all subjects for the random effect analyses. The resulting r map was thresholded to control the false discovery rate (FDR), with an uncorrected voxel level threshold at $p < .001$ and cluster level threshold at $p < .05$.

Model RDMs

We devised two sets of model RDMs, condition-based syntactic complexity models and item-based frequency models. The condition-based syntactic complexity models were three 18 × 18 matrices constructed to assess the neural patterns associated with syntactic complexity (Figure 3). In each model, blue indicates correlated activation patterns due to a shared property (similarity), and red indicates no correlation (dissimilarity). The Complexity model allows us to identify the regions where different degrees of syntactic complexity trigger dissociable processing patterns (Figure 3A, D). This model assumes that each level of syntactic complexity creates a specific and consistent activation pattern (blue), which is dissimilar to the pattern triggered by the other two levels of complexity (red). The Simple versus Hierarchical model (Figure 3B, E) looks for brain regions where simple, nonhierarchical structures (words and short phrases) trigger activation patterns that are similar to each other but different to those for hierarchical structures. Finally, the Word versus Phrase model (Figure 3C, F) distinguishes between the presence and absence of syntactic
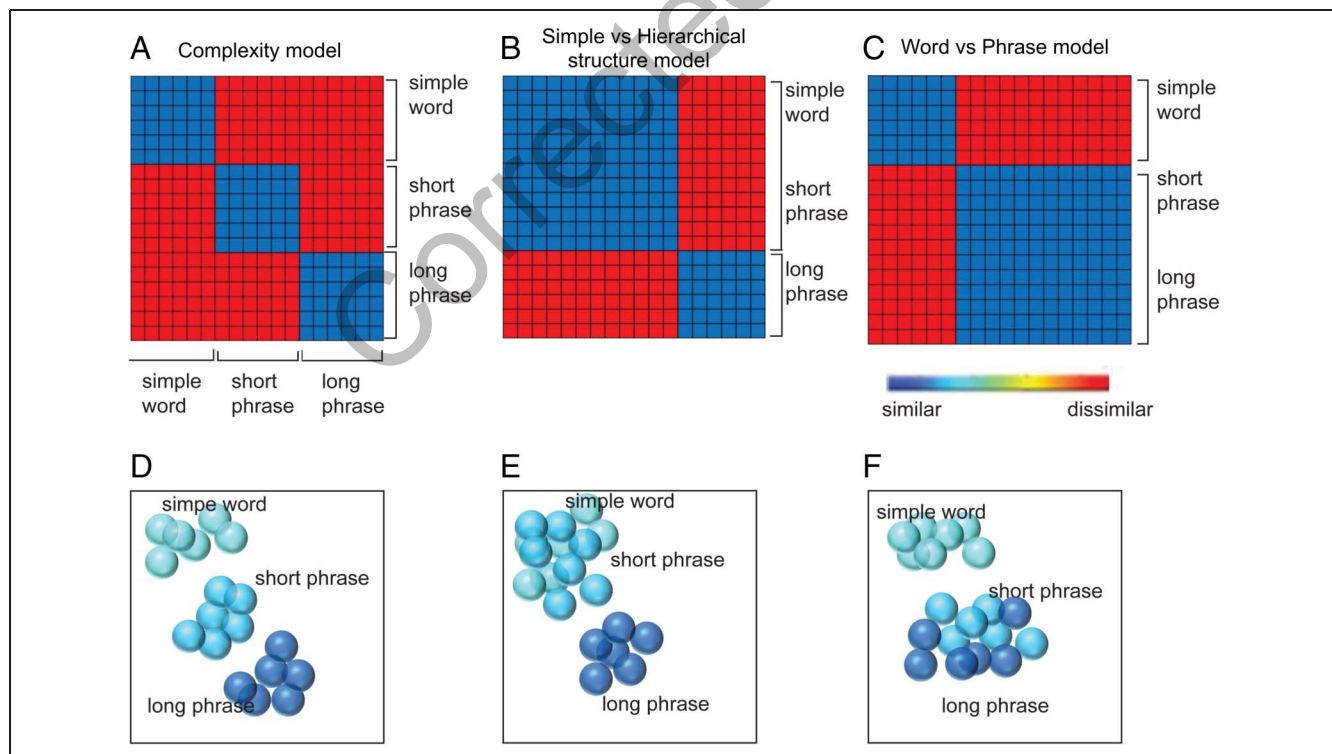

Figure 3. Models used in the RSA syntactic complexity analyses. Top: Syntactic complexity model RDMs for (A) Complexity model, (B) Simple versus Hierarchical structure model, and (C) Word versus Phrase model. Bottom: Correlational structure hypothesized by each model. (D) Complexity model—each level of syntactic complexity (simple word, short phrase, and long phrase) elicits highly correlated spatial patterns but these are distinct from those elicited by the other conditions. (E) Simple versus Hierarchical structure model—simple words and short phrases produce similar patterns which are different from those elicited by long phrases. (F) Word versus Phrase model—the patterns of activity elicited by short and long phrases are highly correlated with each other but are different from those elicited by simple words.
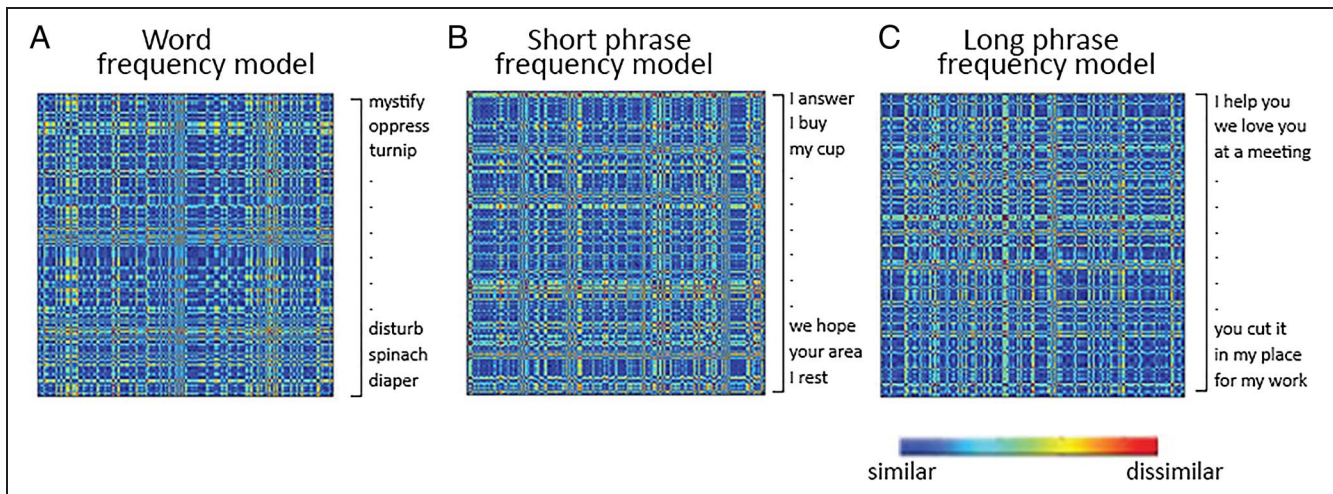
Figure 4. Models used in the RSA frequency analyses: (A) word frequency, (B) short phrase frequency, and (C) long phrase frequency. Each matrix is 180 × 180, where each row or column corresponds to an item used in the experiment. Each model represents the hypothesis that items that are more similar in frequency also have more similar voxel patterns, with the responses expected to cluster together according to their frequency bands (low, mid, high).

complexity by testing for regions where activation patterns for single words differ from activation patterns for phrases (short and long phrases combined).

The item-based frequency models were three models designed to probe the modulation by frequency at each level of syntactic complexity (Figure 4). For each level of syntactic complexity, model RDMs were constructed based on the log frequency of each item, resulting in three 180 × 180 model RDMs for words, short phrases, and long phrases separately (Figure 4A, B, C). These

model RDMs test the hypothesis that items that are more similar in frequency will have more similar activation patterns, whereas items that are more dissimilar in frequency will have more dissimilar activation patterns.

RESULTS

Whole-brain Univariate Analyses

The basic contrast of MuR−silence showed bilateral temporal activation, including STG, MTG, and insular areas,

Figure 5. Univariate effects for simple words. Top: The activation for the contrast "simple word–MuR." Middle: The activation for the contrast "simple verb–MuR." Bottom: The activation for the contrast "simple noun–MuR." All rendered on an inflated canonical brain (voxel level threshold = 0.001; cluster level threshold = 0.05).
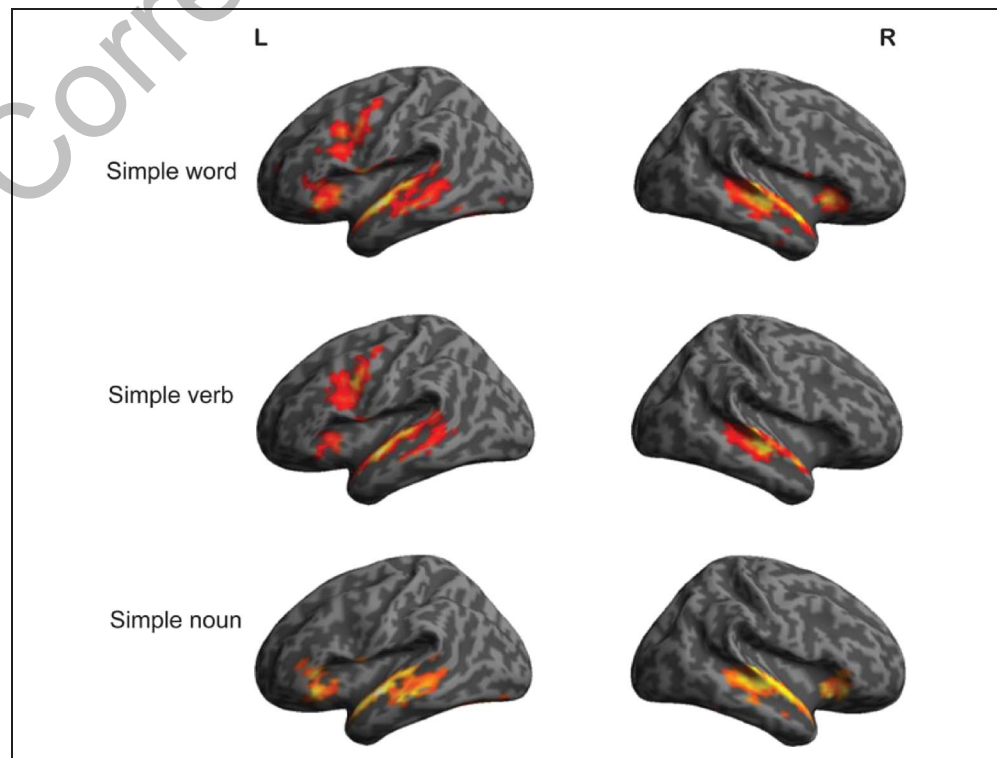
Table 2. The p Values and the MNI Coordinates for Peak Voxels of Significant Clusters for the Contrast "Simple Words–MuR"

Regions	Cluster Level		z Score	Peak Voxel		
	$p_{FDR\text{-}corr}$	Extent		x	y	z
Simple Words–MuR						
L MTG	**<.001**	**3564**	**6.07**	**−64**	**−10**	**−2**
L STG			5.94	−60	−2	−8
L MTG			5.59	−66	−22	0
L superior temporal pole			5.22	−52	1	−12
L IFG (pars orbitalis)			4.6	−38	24	−2
R STG	**<.001**	**3330**	**5.78**	**62**	**−18**	**−2**
R putamen			5.57	30	18	2
R STG			5.25	62	0	−8
L fusiform	**<.001**	**1018**	**4.83**	**−44**	**4**	**36**
L cerebellum			4.57	−48	−58	−28
L ITG			4.39	−40	−30	−14
L precentral	**<.001**	**1086**	**4.77**	**−44**	**4**	**36**
L precentral			3.83	−42	−6	54
L IFG (pars triangularis)			3.25	−38	24	20

Throughout, results were threshold at $p < .001$ voxel level and clusters that survived $p < .05$ corrected for multiple comparisons were considered significant. The highest peaks within a cluster are shown, with the most significant marked in **bold**. MuR = Musical Rain

consistent with earlier research investigating complex auditory processing (cf. Rauschecker & Scott, 2009).

Brain activation related to utterances with different levels of syntactic complexity was assessed by subtracting the acoustic baseline (MuR) from each level of syntactic complexity. Simple words produced clusters in bilateral temporal areas including superior temporal pole, STG, MTG, and ventral IFG (BA 47), and the left precentral gyrus spreading into LIFG BA 45 (Figure 5 and Table 2). A further breakdown of the simple words into verbs and nouns revealed that the activation of L BA 45 was mainly driven by verbs (Figure 5), which activated clusters in bilat-

eral temporal areas and left precentral gyrus and IFG (BA 45). Simple nouns elicited activation in bilateral temporal and ventral IFG regions (BA 47, Figure 5). Short phrases produced bilateral temporal activation primarily in bilateral superior temporal pole and superior and middle temporal areas (Figure 6 and Table 3). The results for the long phrases (Figure 6 and Table 3) showed substantial left IFG activation (BA 45) as well as bilateral temporal activation. Neither short nor long phrases showed significant differences in responses to verbal as opposed to nominal strings

Differences between the three levels of syntactic complexity were tested by directly comparing their activation

Figure 6. Univariate effects for short and long phrases. The activation for the contrast "short phrase–MuR" (top) and "long phrase–MuR" (bottom) rendered on an inflated canonical brain (voxel level threshold = 0.001; cluster level threshold = 0.05).
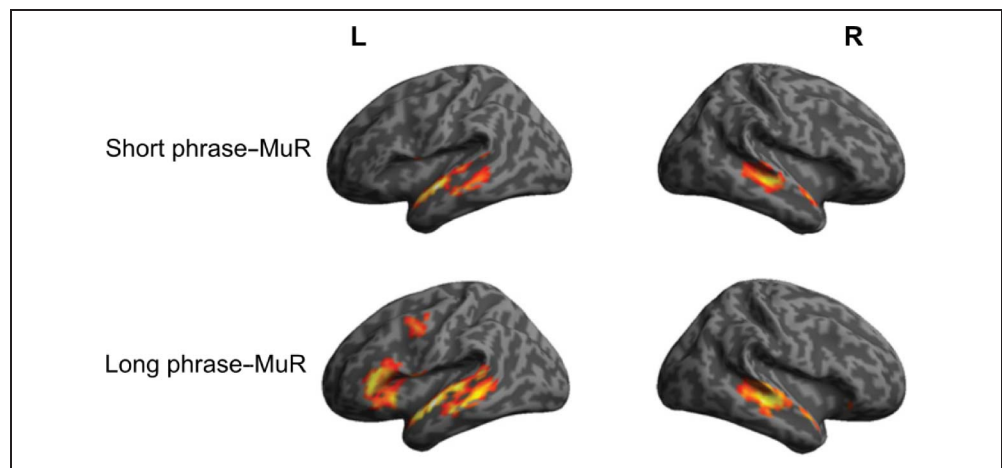
Table 3. The p Values and the MNI Coordinates for Peak Voxels of Significant Clusters for the Contrast "Short Phrases–MuR" and "Long Phrases–MuR"

Regions	Cluster Level			Peak Voxel		
	$p_{FDR\text{-}corr}$	Extent	z Score	x	y	z
Short Phrases–MuR						
L MTG	**<.001**	**1699**	**5.69**	**−60**	**−6**	**−8**
L superior temporal pole			5.49	−54	10	−12
L MTG			4.87	−66	−22	0
R MTG	**<.001**	**960**	**5.01**	**62**	**−20**	**−6**
R MTG			4.9	54	−30	0
R ITG			4.11	42	−30	−6
R superior temporal pole	**.023**	**232**	**4.6**	**56**	**8**	**−14**
R superior temporal pole			4.19	50	16	−18
Long Phrases–MuR						
L superior temporal pole	**<.001**	**4228**	**6.17**	**−56**	**10**	**−10**
L STG			5.8	−60	0	−10
L MTG			5.6	−60	−8	−8
L MTG			5.31	−62	−22	−2
L IFG (pars triangularis)			5.04	−50	28	4
R MTG	**<.001**	**1939**	**5.6**	**54**	**−32**	**0**
R STG			5.3	60	−20	−4
R middle temporal pole			4.78	60	4	−6
L precentral	**.01**	**409**	**4.45**	**−44**	**0**	**44**
R cerebellum	**.01**	**398**	**4.44**	**22**	**−70**	**−44**
R cerebellum			4.28	26	−58	−40
L supplementary motor	**.009**	**467**	**4.13**	**−4**	**14**	**48**
L supplementary motor			3.84	−6	8	58

Figure 7. The activation for the contrast "simple word–short phrase" (top) and "long phrase–short phrase" (bottom) Voxel level threshold = 0.001; cluster level threshold = 0.05.
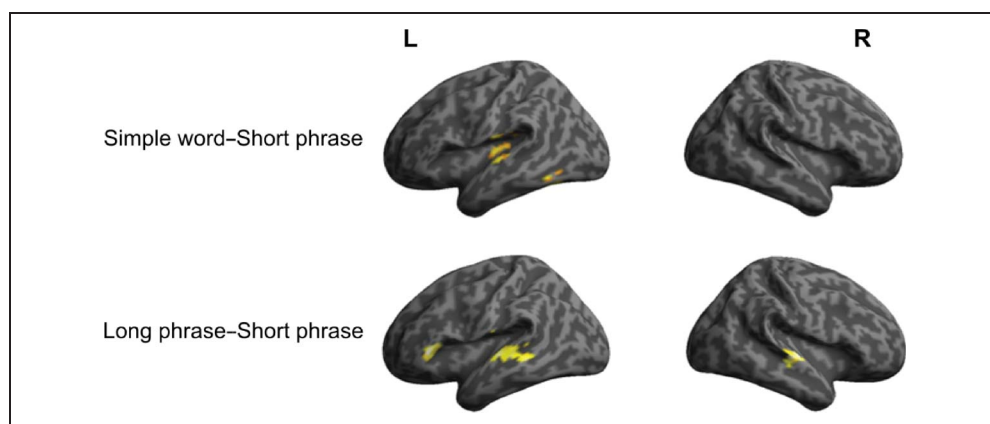
Table 4. The *p* Values and the MNI Coordinates for Peak Voxels of Significant Clusters for the Contrast "Long Phrase–Short Phrase" and "Simple Word–Short Phrase"

Regions	Cluster Level		z Score	Peak Voxel		
	$p_{FDR\text{-}corr}$	Extent		x	y	z
Long Phrases–Short Phrases						
R STG	.008	365	4.34	58	−14	0
R MTG			3.5	66	−28	−2
L IFG (pars orbitalis)	.032	235	4.02	−48	26	−2
L IFG (pars opercularis)			3.54	−54	16	4
L STG	.001	611	4.01	−60	−14	6
L MTG			3.91	−54	−32	0
L MTG			3.82	−56	−22	0
Simple Words–Short Phrases						
L fusiform	.001	580	4.66	−40	−50	−22
L cerebellum			4.35	−48	−58	−28
L STG	.001	568	4.25	−62	−16	4
L medial temporal			3.99	−36	−28	6
L rolandic operculum			3.58	−46	−16	18

after the acoustic baseline (MuR) was subtracted out. Simple words produced more activation than short phrases in left temporal areas (Figure 7 and Table 4). Long phrases showed stronger activation than short phrases in LIFG (BA 47, BA 44) and bilateral temporal areas (Figure 7 and Table 4). No statistically reliable differences were found between simple words and long phrases.

Parametric Modulator Analyses

To assess the effect of frequency, parametric analyses were used in which the log frequency of each word, short phrase, and long phrase was entered as the parametric modulator. This allowed a search for brain regions whose activation amplitude linearly increased or decreased as a function of frequency. The log frequency of simple words negatively correlated with a cluster in left temporal areas, with the peak (at −62 −16 2) located in left MTG. No significant correlations were found with the log frequency of short phrases or long phrases.

Multivariate Whole-brain Searchlight RSA

Condition-based Syntactic Complexity Models

The condition-based syntactic complexity models were constructed to probe similarities and differences between activation patterns associated with different levels of syntactic complexity. Because the clusters found to correlate with some of these models were extensive, covering a set of broadly distributed regions, presenting only a few peak voxels does not adequately represent the results. For example, the first few peak voxels for the Complexity model as reported in Table 5 are in the temporal regions, but the brain images (Figure 8) reveal that LIFG (BA 45) also correlates with the model to a high degree, with almost 50% of the voxels in that region correlating with the model. Therefore, for models that show extensive and widely distributed clusters, further peaks (up to a maximum of 10) are also reported.

The Complexity model, which was constructed to pick up regions that show differentiable activation patterns for simple words, short phrases, and long phrases, correlates with an extensive network of bilateral temporal and left IFG (BA 45). The Simple versus Hierarchical model, which was designed to detect regions where simple sequences pattern together and are dissimilar to hierarchical sequences, correlated with activation in medial and middle temporal regions bilaterally, as well as right middle frontal areas (Figure 8 and Table 5). The Word versus Phrase model, which was designed to search for the areas showing sensitivity to the differences between presence or absence of syntactic complexity, was found to correlate with left IFG (BA 45) and left temporal areas including STG, MTG, and inferior temporal gyrus (ITG; Figure 8 and Table 5).

Table 5. RSA Syntactic Complexity Analyses

| Regions	Cluster Level		z Score	Peak Voxel		
	$p_{FDR-corr}$	Extent		x	y	z
Complexity Model						
L MTG	**<.001**	**103010**	**5.49**	**−51**	**−40**	**10**
L medial temporal			5.29	−33	23	21
L fusiform			5.26	−45	−52	−20
R medial temporal			5.19	27	−7	29
L ITG			5.18	−42	−34	−12
R MTG			5.12	54	−37	−9
L MTG			4.97	−63	−34	−9
L IFG (pars triangularis)* (47%)			3.93	−42	35	3
Simple vs. Hierarchical Model						
R medial temporal	**<.001**	**6615**	**5.92**	**27**	**−16**	**25**
R putamen			5.09	24	14	6
L medial temporal			5.09	−12	−34	29
R mid cingulum			4.67	9	−10	36
R middle frontal			4.65	27	32	29
R cerebellum			4.65	15	−49	−50
R hippocampus			4.51	33	−31	−5
R ITG			4.47	48	−19	−20
L MTG* (25%)			4.38	−51	−64	18
L putamen	**.019**	**120**	**4.05**	**−12**	**11**	**−1**
Word vs. Phrase Model						
L rolandic operculum	**.025**	**176**	**4.15**	**−45**	**20**	**21**
L IFG (pars triangularis)			3.87	−48	29	14
L ITG	**.038**	**130**	**3.99**	**−48**	**−46**	**−16**
L MTG			3.69	−54	−55	6
L MTG			3.61	−66	−43	−5
L STG			3.33	−54	−46	18

p Values and MNI coordinates for the peak voxels of significant clusters correlated with the Complexity model, the Simple vs. Hierarchical model, and the Word vs. Phrase model (voxel level threshold = 0.001; cluster level threshold = 0.05). Regions marked with an asterisk (*) are ROIs that are included in the local maxima of the current cluster. The percentage indicates the percentage of the voxels in the region found in the cluster.

Item-based Frequency Models

Three item-based RSA models were built to look for neural activity patterns that are modulated by simple word frequency, short phrase frequency, and long phrase frequency. The simple word frequency model correlated primarily with left frontal regions in superior and middle frontal cortex and left BA 47. Right hemisphere model fit included right inferior occipital cortex and a cluster with the peak in the hippocampus, extending to right superior temporal areas (Figure 9 and Table 6). No brain regions were found to correlate with the short phrase frequency or the long phrase frequency model.
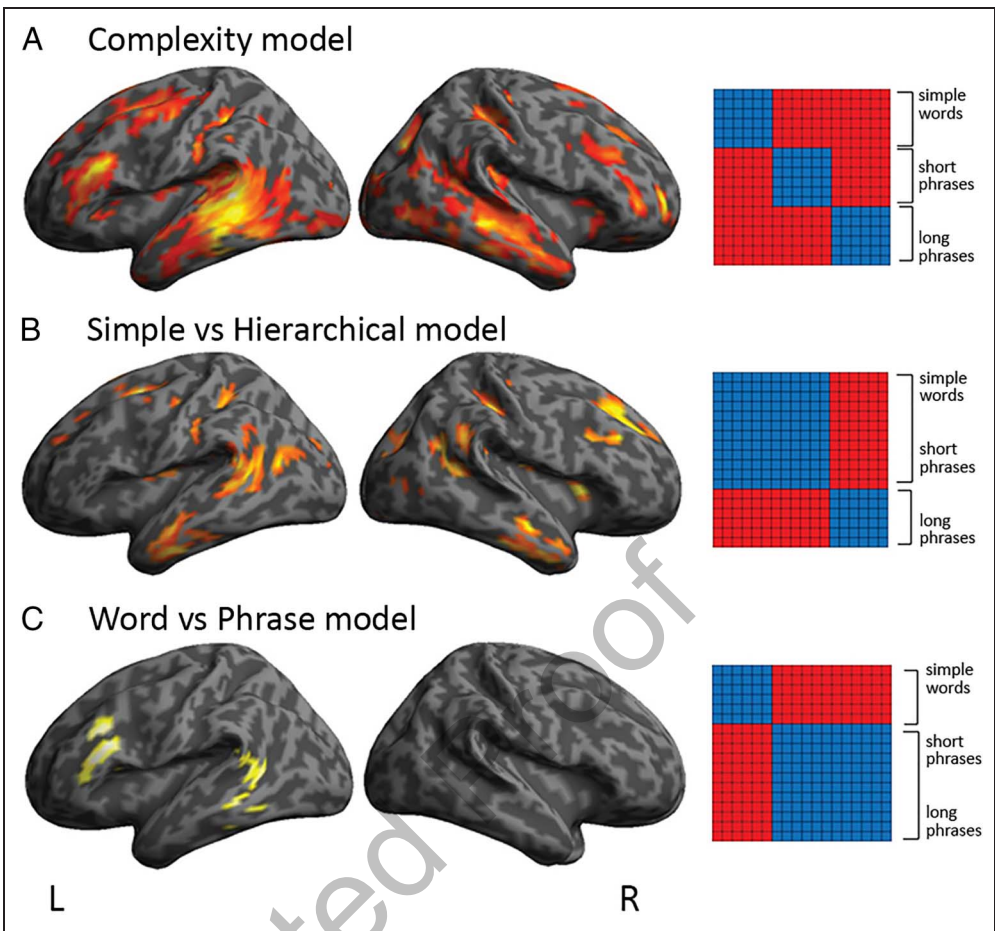
DISCUSSION

The goal of this fMRI study was to evaluate a range of hypotheses about how the complexity and frequency of spoken words and phrases influence the way they are processed and represented by the language networks. A particular focus was on the role of frequency, asking whether different linguistic strings ranging from single words to long phrases were stored and retrieved during language processing as a function of their frequency, in line with the suggestions made by distributional approaches to grammar (Bybee & McClelland, 2005). Within the framework of the dual neurobiological language system adopted here (Marslen-Wilson, Bozic, & Tyler, 2014), this would suggest that the more frequent words and phrases could be stored and retrieved by the tempo-

ral bihemispheric processing network, whereas the primary role of the left frontotemporal network would be to analyze more complex and infrequent strings lacking memory traces.

To this end, we manipulated two variables: syntactic complexity (three levels: low, simple words; mid, short phrases; high, longer phrases) and frequency (three levels: low, mid, and high). A combination of standard univariate methods and multivariate RSA was used to investigate how these variables affect the overall activation amplitudes and whether these activations actually encode comparable processes. We begin with an overview of the differential effects of the syntactic complexity contrasts, because these provide an essential framework for evaluating potential modulatory effects of frequency.

Table 6. RSA Word Frequency Results

Regions	Cluster Level			Peak Voxel		
	$p_{FDR\text{-}corr}$	Extent	z Score	x	y	z
Word Frequency Model						
L cerebellum	**.007**	**153**	**5.06**	**−9**	**−79**	**−28**
R vermis			4.19	3	−58	−16
R vermis			4.14	0	−73	−20
L medial temporal	**.001**	**227**	**4.64**	**−30**	**41**	**10**
L IFG (pars orbitalis)			4.18	−42	35	−9
L orbital frontal			3.99	−21	29	−9
L superior frontal			3.67	−24	56	21
L cerebellum	**.038**	**91**	**4.33**	**−30**	**−58**	**−24**
R fusiform	**.046**	**81**	**4.13**	**36**	**−73**	**−16**
R inferior occipital			3.6	45	−73	−1
R hippocampus	**.019**	**115**	**3.84**	**33**	**−37**	**10**
R hippocampus			3.64	33	−28	33

The *p* values and the MNI coordinates for peak voxels of significant clusters correlated with the item-based word frequency model (voxel level threshold = 0.001; cluster level threshold = 0.05).

Distribution of Syntactic Complexity Effects

We first used classic univariate methods to establish how simple words, short phrases, and long phrases engage the underlying neural architecture. Simple words elicited activation in bilateral temporal and ventral IFG areas and in L BA45. Because this LIFG activation was not predicted, additional analyses were performed to investigate this result further. Looking at the activation for simple verbs and nouns separately, it emerged that the L BA45 activation was triggered primarily by verbs, which is consistent with previous findings of increased left frontotemporal engagement for verbs over nouns due to the complex argument structures of verbs (e.g., Shapiro et al., 2005; Tyler, Bright, Fletcher, & Stamatakis, 2004). Simple nouns, in contrast, elicited activation only in bilateral temporal and ventral IFG regions (BA 47). Although the simple nouns used in this experiment were words without inflectional affixes, they were typically multisyllabic words with onset-embedded pseudostems (e.g., *cutlet, turnip*), which are likely to generate cohort competition during spoken language comprehension. Our results are consistent with studies showing that bilateral temporal areas are implicated in the accessing of semantic representations of words (Binder, Desai, Graves, & Conant, 2009; Jung-Beeman, 2005) as well as studies showing that bilateral IFG and temporal areas are engaged in processing perceptually complex words (e.g.,

Carota, Bozic, & Marslen-Wilson, 2016; Zhuang, Tyler, Randall, Stamatakis, & Marslen-Wilson, 2014; Bozic et al., 2010).

Turning to the phrasal conditions, short phrases engaged bilateral superior and middle temporal regions, including anterior to posterior STG and middle to posterior MTG, but with no evidence for selective LIFG activation. Short phrases, which involve a minimum level of syntactic combinatorial processing, have received little attention in the neuroimaging literature: exceptions include Bemis and Pylkkanen (2011) and Bozic et al. (2015). Bemis and Pylkkanen (2011) found that basic syntactic operations of linking adjectives and nouns activate left anterior temporal areas, whereas Bozic et al. (2015) found that short phrases involving pronoun and verb or determiner and noun combinations correlate with bilateral anterior and posterior temporal regions. The findings here for short phrases are in line with the previous Bozic et al. results and are consistent with the view that minimal phrase structures, capturing linear adjacency relationships, engage mainly temporal regions and do not necessarily involve left IFG regions.[1] However, given that these temporal areas are also associated with the basic mapping between sound and lexico-semantic representations (Hickok & Poeppel, 2007), it is unclear, without taking frequency into account, whether the engagement of temporal areas for short phrases reflects basic combinatorial structure building or retrieval of whole phrases, a question we explore below.

For the syntactically most complex condition, long phrases, there was prominent activation in left IFG, especially BA 45, as well as in bilateral STG and MTG. This finding is in line with previous studies demonstrating the important role of the LIFG, especially the dorsal portion (BA 44/45), in hierarchical structure processing—though much of this evidence comes from studies using complex combinations such as scrambled sentences (Friederici et al., 2006) and center-embedded sentences (Makuuchi et al., 2009). The results here show that dorsal LIFG can also be activated by syntactic structures that are minimally hierarchically complex and which lack nonlocal dependencies. The involvement of the LIFG in minimal hierarchical structure processing is further strengthened by the finding that long phrases as a whole elicited more activation than short phrases in LIFG and left posterior temporal regions. This network, linking left BA 44/45 and left STG and MTG has been widely claimed to support linguistic combinatorial processing (Friederici, 2011; Tyler & Marslen-Wilson, 2008; for a recent meta-analysis of this literature, see Hagoort & Indefrey, 2014). The elevated activation seen here of this network arguably reflects the extra syntactic computation demands imposed by long phrases over short phrases.

The results of the whole-brain searchlight RSA allowed us to test the specific processes that underlie the distribution of activations observed in the univariate analyses. The Complexity model, which dissociates between all three types of syntactic complexity (no syntactic complexity, simple combinatorial structure, and hierarchical combinatorial structure), is found to correlate with an extensive network of bilateral frontotemporal regions. This result shows that, even with the three types of sequences triggering similar activation amplitudes in this processing network, this is likely to reflect averages across different types of processes, to which univariate analyses are not sensitive. However, this model by itself cannot illuminate the links between regions and specific types of process or computation.

We turned to the other two RSA models to test the specific computational properties of the brain regions involved. The Simple versus Hierarchical model groups together simple words and short phrases, on the one hand, and long phrases, on the other hand. This model RDM is designed to detect the activation signature of simple comprehension processes that the first two conditions have in common and to dissociate these from the processing of hierarchical structures in the long phrases. This model primarily correlates with bilateral temporal areas, implying that the spatial patterning of neural activity in these regions differentiates between the processing of items without hierarchical structure (words and short phrases) and items that do have such structure (long phrases). These results are in line with the univariate results reported above, which show shared bilateral temporal activation for simple words and short phrases, as well as recent findings that bilateral temporal regions provide

a basis for lexical interpretation of spoken utterances, both at the level of simple words and minimal phrases (Bozic et al., 2015; Bemis & Pylkkanen, 2013).

The Word versus Phrase model, which was designed to look for brain regions that encode differences in activation patterns due to the presence or absence of syntactic complexity, correlates specifically to activity patterns in left BA 45 and left posterior temporal regions. Consistent with the univariate results, this implies that the left frontotemporal network is critically sensitive to the absence or presence of syntactic complexity and combinatorial processing in particular. More generally, these results fit well with findings from the large body of studies linking left BA 44/45 and left temporal activation with combinatorial syntactic processing (e.g., Tyler et al., 2011; Makuuchi et al., 2009; Friederici et al., 2006; Embick, Marantz, Miyashita, O'Neil, & Sakai, 2000).

In summary, the univariate and RSA results show that the bilateral and the left-lateralized language networks encode complementary aspects of syntactic processing. As in the previous Bozic et al. (2015) research, bilateral temporal areas were engaged by the processing of short phrases, with the RSA results further specifying that activity patterns in this network differentiate between syntactically simpler utterances and those that require more complex processing of hierarchical structures. The LH frontotemporal network was shown to be sensitive to the presence or absence of syntactic complexity in the RSA analyses, with the univariate results showing that selective engagement of this network requires the additional hierarchical complexity present in the long phrases.

The Effects of Frequency

The complementary goal of this study was to investigate the effects of frequency at different levels of syntactic complexity. Given the results summarized in the previous section, we would expect to see such effects, if present, affecting bilateral temporal activity for the short phrases, with additional effects in LIFG for the long phrases.

Where simple words are concerned, the parametric modulator analyses showed that the frequency of simple words negatively correlated with activation in left temporal areas, mainly in the middle portion of the left MTG—a region commonly implicated in the retrieval of stored lexical information (Price, 2012). The direction of the frequency effects is consistent with earlier research, which has also found stronger activation for lower frequency words (e.g., Nakic et al., 2006; Fiebach, Friederici, Müller, & von Cramon, 2002). This implies that the less frequent a word is, the more effortful it is to access and retrieve the word, resulting in higher levels of neural activation.

For short phrases, the storage hypothesis predicted that frequent phrases were more likely to be stored as whole forms, so that the activation they trigger would decrease as a function of frequency. The univariate findings did not support this hypothesis: There were no regions

in which activation correlated with frequency of short phrases, providing no evidence that short phrases are stored and retrieved as whole forms during spoken language comprehension. This, in turn, suggests that the bilateral activation for short phrases found here and in previous studies (e.g., Bozic et al., 2015) is not because these simple phrases are being retrieved as if they were stored like simple words. Despite behavioral studies showing sensitivity to multiword sequences during reading (Arnon & Snider, 2010; Tremblay & Baayen, 2010), as well as the predictions of distributional approaches to grammar (Bybee & McClelland, 2005), we see no neural evidence that simple two-word phrases are retrieved as a whole during spoken language comprehension. Instead the bilateral temporal activation found here is consistent with the proposal that these temporal regions can support simple combinatorial processes (Bozic et al., 2015) and is therefore inconsistent with the obligatory syntax view that locates all such processes in dorsal LIFG.

The related prediction for long phrases was that frequency variation would modulate the amount of activation in the core left frontotemporal network, including dorsal LIFG (BA 44 and 45). To the contrary, no regions were detected, that significantly correlated with the frequency of the hierarchically complex long phrases, providing no evidence that such sequences are stored and retrieved during spoken language comprehension. This absence of left IFG modulation by frequency suggests that, inconsistent with the key hypothesis of a storage-based view, the core LH network for hierarchical syntactic structure processing is equally engaged by long phrases irrespective of their frequency.

The spatial patterns associated with frequency for the three levels of syntactic complexity were also explored using item-based RSA analyses. The analyses showed that word frequency correlates with patterns in frontal regions and in left BA 47. The ventral part of the left IFG plays an important part in lexicosemantic processing and is particularly important for lexical competition and selection processes (e.g., Zhuang et al., 2014). This implies that the words that are similar in frequency induce similar lexical retrieval competition load in the anterior part of the LIFG. Item-based RSA analyses again failed to find patterns correlating with short phrase or long phrase frequency. Echoing the results of the univariate analyses, item-based searchlight RSA did not support the storage hypothesis stating that short or long phrases are stored and retrieved during spoken language comprehension. Instead, RSA show that, in line with the obligatory syntactic parsing account, though simple words are stored, there is no evidence that short and long phrases are retrieved as whole forms during language comprehension, implying that syntactic parsing is an obligatory process for the two phrasal conditions.

Taken as a whole, the findings in this experiment are strongly inconsistent with predictions following from the storage account and show that distributional frequency of syntactically complex items plays little role during spoken language comprehension—specifically that they are not retrieved as stored forms during language processing. Instead our findings favor an obligatory syntactic parsing account applying to syntactically complex items across all frequency levels, in which hierarchical syntactic structure is the main factor driving activation in the left frontotemporal network. Finally, taking the absence of frequency effects for phrases here together with the bilateral temporal activation observed here and in Bozic et al. (2015), we conclude that the neural activity in bilateral temporal areas cannot be explained in terms of retrieval of whole phrases, and these areas in fact have some degree of combinatorial processing capacity.

Notes

1. Also consistent with Bozic et al (2015), we saw no differences between the activations for short verb and noun phrase. To account for these results we suggest that, in the context of unfolding acoustic information, hearing a pronoun or an article is deterministic with respect to the grammatical properties of the subsequent element. This renders the combinatorial and representational differences between verbs and nouns less relevant, revealing instead the common underlying mechanism of simple conctituent structure grouping. This activates the bilateral circuit that supports the linear groupings between adjacent elements.

REFERENCES

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62,* 67–82.

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

Bemis, D. K., & Pylkkanen, L. (2013). Basic linguistic composition recruits the left anterior temporal lobe and left angular gyrus during both listening and reading. *Cerebral Cortex, 23,* 1859–1873.

Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex, 19,* 2767–2796.

Bozic, M., Fonteneau, E., Su, L., & Marslen-Wilson, W. D. (2015). Grammatical analysis as a distributed neurobiological function. *Human Brain Mapping, 36,* 1190–1201.

Bozic, M., Tyler, L. K., Ives, D. T., Randall, B., & Marslen-Wilson, W. D. (2010). Bihemispheric foundations for human speech comprehension. *Proceedings of the National Academy of Sciences, U.S.A., 107,* 17439–17444.

Bybee, J. L., & McClelland, J. L. (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *Linguistic Review, 22,* 381–410.

Carota, F., Bozic, M., & Marslen-Wilson, W. (2016). Decompositional representation of morphological compexity: Multivariate fMRI evidence from italian. *Journal of Cognitive Neuroscience, 28,* 1878–1896.

Chee, M. W. L., Westphal, C., Goh, J., Graham, S., & Song, A. W. (2003). Word frequency and subsequent memory effects studied using event-related fMRI. *Neuroimage, 20,* 1042–1051.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge, MA: MIT Press.

Cusack, R., Vicente-Grabovetsky, A., Mitchell, D. J., Wild, C. J., Auer, T., & Linke, A. C., , et al. (2015). Automatic analysis (aa): Efficient neuroimaging workflows and parallel processing using Matlab and XML. *Frontiers in Neuroinformatics, 8,* 90.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology, 42,* 317–367.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., & Sakai, K. L. (2000). A syntactic specialization for Broca's area. *Proceedings of the National Academy of Sciences, U.S.A., 97,* 6150.

Fiebach, C. J., Friederici, A. D., Müller, K., & von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of Cognitive Neuroscience, 14,* 11–23.

Friederici, A. D. (2011). The brain basis of language processing: From structure and function. *Physiological Reviews, 91,* 1357–1392.

Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. D., & von Cramon, D. Y. (2006). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cerebral Cortex, 16,* 1709–1717.

Friederici, A. D., Wang, Y., Herrmann, C. S., Maess, B., & Oertel, U. (2000). Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. *Human Brain Mapping, 11,* 1–11.

Griffiths, J. D., Marslen-Wilson, W. D., Stamatakis, E. A., & Tyler, L. K. (2013). Functional organization of the neural language system: Dorsal and ventral pathways are critical for syntax. *Cerebral Cortex, 23,* 139–147.

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

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Frontiers in Psychology, 4,* 416.

Hagoort, P. (2014). Nodes and networks in the neural architecture for language: Broca's region and beyond. *Current Opinion in Neurobiology, 28,* 136–141.

Hagoort, P., & Indefrey, P. (2014). The neurobiology of language beyond single words. *Annual Review of Neuroscience, 37,* 347–362.

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

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

Janssen, N., & Barber, H. A. (2012). Phrase frequency effects in language production. *PLoS One, 7,* e33202.

Jung-Beeman, M. (2005). Bilateral brain processes for comprehending natural language. *Trends in Cognitive Sciences, 9,* 512–518.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 3863–3868.

Makuuchi, M., Bahlmann, J., Anwander, A., & Friederici, A. D. (2009). Segregating the core computational faculty of human language from working memory. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 8362–8367.

Marslen-Wilson, W. D., Bozic, M., & Tyler, L. K. (2014). Morphological systems in their neurobiological contexts. In M. S. Gazzaniga & G. R. Mangun (Eds.), *The cognitive neurosciences (5th ed.).* Cambridge, MA: MIT Press.

Nakic, M., Smith, B. W., Busis, S., Vythilingam, M., & Blair, R. J. (2006). The impact of affect and frequency on lexical decision: The role of the amygdala and inferior frontal cortex. *Neuroimage, 31,* 1752–1761.

Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLoS Computational Biology, 10,* 1–11.

Pinker, S. (1999). *Words and rules.* New York, NY: Harper Perennial.

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage, 62,* 816–847.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience, 12,* 718–724.

Rolheiser, T., Stamatakis, E. A., & Tyler, L. K. (2011). Dynamic processing in the human language system: Synergy between the arcuate fascicle and extreme capsule. *Journal of Neuroscience, 31,* 16949–16957.

Shapiro, K. A., Mottaghy, F. M., Schiller, N. O., Poeppel, T. D., Fluss, M. O., Muller, H. W., et al. (2005). Dissociating neural correlates for nouns and verbs. *Neuroimage, 24,* 1058–1067.

Tremblay, A., & Baayen, H. (2010). Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 151–173). London: Continuum International.

Tyler, L. K., Bright, P., Fletcher, P., & Stamatakis, E. A. (2004). Neural processing of nouns and verbs: The role of inflectional morphology. *Neuropsychologia, 42,* 512–523.

Tyler, L. K., & Marslen-Wilson, W. (2008). Frontotemporal brain systems supporting spoken language comprehension. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences, 363,* 1037–1054.

Tyler, L. K., Marslen-Wilson, W. D., Randall, B., Wright, P., Devereux, B. J., Zhuang, J., et al. (2011). Left inferior frontal cortex and syntax: Function, structure and behaviour in left-hemisphere damaged patients. *Brain, 134,* 415–431.

Tyler, L. K., Shafto, M. A., Randall, B., Wright, P., Marslen-Wilson, W. D., & Stamatakis, E. A. (2010). Preserving syntactic processing across the adult life span: The modulation of the frontotemporal language system in the context of age-related atrophy. *Cerebral Cortex, 20,* 352–364.

Uppenkamp, S., Johnsrude, I. S., Norris, D., Marslen-Wilson, W., & Patterson, R. D. (2006). Locating the initial stages of speech-sound processing in human temporal cortex. *Neuroimage, 31,* 1284–1296.

Wright, P., Randall, B., Marslen-Wilson, W. D., & Tyler, L. K. (2011). Dissociating linguistic and task-related activity in the left inferior frontal gyrus. *Journal of Cognitive Neuroscience, 23,* 404–413.

Wright, P., Stamatakis, E. A., & Tyler, L. K. (2012). Differentiating hemispheric contributions to syntax and semantics in patients with left-hemisphere lesions. *Journal of Neuroscience, 32,* 8149–8157.