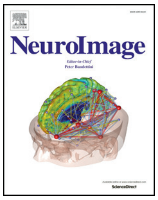
NeuroImage 132 (2016) 293–300

Contents lists available at ScienceDirect

NeuroImage

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



Language structure in the brain: A fixation-related fMRI study of syntactic surprisal in reading

John M. Henderson *, Wonil Choi, Matthew W. Lowder, Fernanda Ferreira

University of California, Davis, USA

ARTICLE INFO

Article history:
Received 31 August 2015
Accepted 15 February 2016
Available online 22 February 2016

Keywords:
Language
Syntax
Surprisal
Reading
Eye movements
fMRI

ABSTRACT

How is syntactic analysis implemented by the human brain during language comprehension? The current study combined methods from computational linguistics, eyetracking, and fMRI to address this question. Subjects read passages of text presented as paragraphs while their eye movements were recorded in an MRI scanner. We parsed the text using a probabilistic context-free grammar to isolate syntactic difficulty. Syntactic difficulty was quantified as syntactic surprisal, which is related to the expectedness of a given word's syntactic category given its preceding context. We compared words with high and low syntactic surprisal values that were equated for length, frequency, and lexical surprisal, and used fixation-related (FIRE) fMRI to measure neural activity associated with syntactic surprisal for each fixated word. We observed greater neural activity for high than low syntactic surprisal in two predicted cortical regions previously identified with syntax: left inferior frontal gyrus (IFG) and less robustly, left anterior superior temporal lobe (ATL). These results support the hypothesis that left IFG and ATL play a central role in syntactic analysis during language comprehension. More generally, the results suggest a broader cortical network associated with syntactic prediction that includes increased activity in bilateral IFG and insula, as well as fusiform and right lingual gyri.

Introduction

A complete theory of the human brain must include a description of the neural networks responsible for language processing. Comprehension of connected sentences requires the ability to retrieve words from the lexicon and relate those words to each other, as the input is encountered in real time. The system for creating these groupings is known as syntax, and is generally thought of as a set of computations specifying syntactic categories such as noun and verb, and stating how those categories combine into increasingly larger constituents such as noun phrases, verb phrases, and clauses. The constituent structure of a sentence is the frame on which interpretations are built; the same words will have radically different meanings depending on that structure (e.g., *the dog bit the man* vs. *the man bit the dog*). Thus, the human ability to understand language is based both on lexical knowledge and syntactic computations.

In this study, we focused on understanding how the human brain implements the syntactic component of the human language faculty. Specifically, we were interested in the nature of the cortical systems that compute syntactic representations during online language comprehension. Three regions have traditionally been identified as candidates for syntactic processing (Fedorenko et al., 2012; Friederici and Kotz, 2003; Grodzinsky and Friederici, 2006; Kaan and Swaab, 2002). Historically, Broca's area or left inferior frontal gyrus (IFG) has been most associated with syntax (Ben-Shachar et al., 2003; Ben-Shachar et al., 2004; Bornkessel-Schlesewsky et al., 2009; Caplan et al., 2008; Caplan et al., 2000; Carramazza and Zurif, 1976; Dapretto and Bookheimer, 1999; Embick et al., 2000; Friederici et al., 2006; Grodzinsky, 2001; Just et al., 1996; Santi and Grodzinsky, 2007; Stromswold et al., 1996). In addition, the left anterior temporal lobe (ATL) has often been linked to computations related to the manipulation of syntactic structure (Brennan et al., 2010; Dronkers et al., 2004; Friederici et al., 2000a; Friederici et al., 2003; Friederici et al., 2000b; Humphries et al., 2005; Humphries et al., 2006; Mazoyer et al., 1993; Noppeney and Price, 2004; Rogalsky and Hickok, 2009; Tyler et al., 2011; Vandenberghe et al., 2002). Finally, several studies suggest a potential role for left posterior superior temporal sulcus and gyrus (STS/STG) in syntactic processing (Constable et al., 2004; Cooke et al., 2002; Hasson et al., 2006; Tyler et al., 2011).

A challenge for research investigating the neural foundations of syntax has been finding an appropriate method for varying syntactic difficulty and measuring its neural effects. Typically this has meant (a) manipulating syntactic complexity within sentences, using (for example) garden-path sentences, subject versus object relative clauses, or active versus passive sentences, and (b) using a secondary task to ensure that participants are processing the sentence, such as sentence memorization or judgments of grammaticality or acceptability. This approach is useful, but sentence-level manipulations mean that only a

---

* Corresponding author at: Center for Mind and Brain, 267 Cousteau Place, University of California, Davis, CA 95618, USA.

E-mail address: johnhenderson@ucdavis.edu (J.M. Henderson).

http://dx.doi.org/10.1016/j.neuroimage.2016.02.050
1053-8119/

single data point can be generated for each sentence, limiting statistical power and generalizability. Furthermore, different types of manipulations may engage different cognitive systems and lead to differences in results. Similarly, the use of a secondary task can induce specific processing strategies that may account for differences in effects across studies, and findings may reflect activity arising from a mixture of linguistic processing and processing related to memorizing, evaluating grammaticality, and so on. An alternative approach would be to generate a syntactic difficulty measure for every word, and to take neuroimaging measurements of processing based on this measure for each word in a manner that does not require a secondary task. This is the approach we adopted in the present study.

Recent research in computational linguistics has been aimed at developing metrics to objectively quantify word-by-word language difficulty. In our study, we capitalized on this work to quantify syntactic difficulty using a metric called surprisal (Hale, 2001; Levy, 2008). Surprisal is an information-theoretic concept that reflects the expectedness of each word given its preceding context. Surprisal is an increasingly influential concept in the language sciences because it reflects moment-by-moment processing operations related to the system's attempts to connect the current input to the left context, based on the comprehender's history of language use (Boston et al., 2008; Demberg and Keller, 2008; Hale, 2001; Levy, 2008). Unlike some other measures of text complexity, a surprisal value can be generated for every word of a text, making it methodologically feasible to relate syntactic difficulty to brain activity on a word-by-word basis during the reading of connected texts. The surprisal value of word w_i is defined as the negative log probability of w_i, given the words that have come before it in the sentence (i.e., $w_1, w_2 \ldots w_{i-1}$).

$$\text{Surprisal}(w_i) = -\log P(w_i | w_1 \ldots w_{i-1}).$$

Higher surprisal values are associated with greater processing difficulty: Surprisal values reliably predict word-by-word reading times derived from eye movement measures (e.g., Boston et al., 2008; Demberg and Keller, 2008; Demberg et al., 2013; Levy, 2008), the N400 event-related potential (ERP) component (Frank et al., 2015), and fMRI activation (Hale et al., 2015; Willems et al., in press).

Importantly, surprisal can be generated at a number of levels of representation. Roark et al. developed an incremental top-down parser (Roark, 2001; Roark et al., 2009) that separates total surprisal into lexical surprisal and syntactic surprisal using an algorithm that builds sets of partial derivations and weights them according to a probabilistic context-free grammar (PCFG). In this parser, total surprisal is decomposed into portions associated with building lexical terminal items (i.e., words) in the parse tree (lexical surprisal) and syntactic structure associated with building nonterminal syntactic nodes (syntactic surprisal). In an empirical validation of these metrics, Roark et al. (2009) showed that self-paced reading times were more accurately predicted from separate estimates of lexical and syntactic surprisal than from a surprisal metric that combined them. Importantly for our purpose of isolating syntactic difficulty, syntactic surprisal in the Roark et al. study was uncorrelated with both lexical surprisal and lexical frequency.

Our central aim was to investigate syntactic processes using syntactic surprisal as a tool, but the present study also provided us with an opportunity to investigate neural processes related to surprisal and prediction more generally. Recent interest in surprisal during language processing bears on the roles of constraint and predictability in language processing and cognition. Evidence has been accumulating that the brain should be viewed as a “prediction machine” (Clark, 2013) that anticipates what will happen next (Bar, 2009; Den Ouden et al., 2012; Friston, 2010; Lupyan and Clark, 2015; Rao and Ballard, 1999). This approach has been productively applied to the field of language processing, with recent findings indicating that the efficiency of language processing emerges in part from the use of prediction by adults and even children, and during both reading and listening (e.g., Altmann and Kamide, 1999; Kamide et al., 2003; Kutas et al., 2011; Mani and Huettig, 2012; Misyak et al., 2010; Garrod, 2004; Pickering and Garrod, 2013; Smith and Levy, 2013; Van Berkum et al., 2005; Wicha et al., 2003; see Kuperberg and Jaeger, 2016; Staub, 2015, for recent reviews). By conducting a whole-brain analysis of functional activation as a function of syntactic surprisal, the present study provided the opportunity to begin to study predictive processing for syntax using functional magnetic resonance imaging (fMRI) in the context of natural reading.

In our study, subjects read natural stories (e.g., The Emperor's New Clothes) presented as paragraphs in an MRI scanner while both eye movements and BOLD signals were recorded. Because effects in an fMRI study may be influenced by inclusion of a secondary task, our general approach is to understand complex cognitive processes in a natural context, an approach that has been used with auditory narratives (Brennan et al., 2010; Wilson et al., 2008), natural reading (Altmann et al., 2014; Choi et al., 2014; Choi and Henderson, 2015), film viewing (Bartels and Zeki, 2004; Hasson et al., 2004; Nishimoto et al., 2011), and active photograph viewing (Choi and Henderson, 2015; Henderson and Choi, 2015; Stansbury et al., 2013). Here we asked participants simply to read naturally for comprehension. To investigate neural activity associated with word-by-word syntactic difficulty, we used fixation-related (FIRE) fMRI, a technique that combines eyetracking with the BOLD response to examine neural activity as a function of the currently fixated item (Henderson and Choi, 2015; Henderson et al., 2015; Richlan et al., 2014; Marsman et al., 2012).

The benefits of the FIRE fMRI approach arise from two key features: First, participants perform no secondary task; instead, they simply read connected text as they normally would, for comprehension. Any observed effects therefore cannot be attributed to task-specific processes. Second, because participants read connected texts, each fixated word provides a potential data point, resulting in large amounts of data for each subject. The result is greater statistical power. Moreover, based on our previous work showing that readers' eye movements while in an MRI scanner are similar to those observed outside the scanner (Choi and Henderson, 2015; Choi et al., 2014; Henderson et al., 2015), the eye movement data obtained using FIRE fMRI allow us to verify that subjects are reading normally for meaning even though they are not required to perform any sort of separate comprehension task.

In sum, syntactic difficulty was operationalized word-by-word by syntactic surprisal. Neural activity was measured for each fixated word using co-registration of eyetracking and fMRI. We hypothesized a priori that syntactic computations are supported by three main cortical regions: left IFG, left ATL, and left posterior STS/STG, but we also examined the general whole-brain activation pattern associated with syntactic surprisal.

Method

Participants

Forty-three right-handed native speakers of English between the ages of 18 and 35 were recruited from the University of South Carolina student population. Three did not finish the experiment, leaving data from 40 participants (13 male) for inclusion in the analyses (mean age: 21.35). All participants reported normal or corrected-to-normal vision, gave informed consent, were screened for MRI safety, and received either $20 or course credit for their participation. The study protocol was approved by the University of South Carolina Institutional Review Board.

Materials

Each participant read the same 22 paragraphs, 11 modified from the short story *The Emperor's New Clothes* by Hans Christian Andersen, and

11 modified from a practice test for the Nelson Denny reading assessment (see Supplementary materials). Paragraphs ranged in length from 49 to 66 words.

Language statistics

Lexical frequency counts and surprisal values were obtained for all words in the 22 paragraphs. Log-transformed lexical frequency counts were generated from the SUBTLEXus corpus (Brysbaert and New, 2009). Each sentence was parsed using Roark's (Roark, 2001; Roark et al., 2009) incremental top-down PCFG parser trained on the Wall Street Journal corpus of the Penn Treebank (Marcus et al., 1993) to generate syntactic surprisal values for each word.

Given our theoretical question, it was important to generate a syntactic surprisal value for each word over a set of words that was as much as possible independent of other lexical and contextual influences. Over the entire set, words at the extremes of the syntactic surprisal distribution differed in frequency and length across high- and low-surprisal values. We therefore selected all of the words in the central two surprisal quartiles for analysis. To generate high- and low-surprisal conditions for the fMRI analysis, we then divided the selected words by median-split to form two syntactic surprisal conditions. As shown in Table 1, the two conditions were well matched on the nuisance variables of lexical surprisal, word frequency, and word length. Because surprisal values did not differ by position, the syntactic surprisal values included in the analysis were representative of all word positions in the sentences. The correlations between continuous values of syntactic surprisal and the nuisance variables were also low and nonsignificant: lexical surprisal $r = .014$, word frequency $r = .008$, and word length $r = .002$.

Apparatus

Stimuli were presented using an Avotec Silent Vision 6011 projector in its native resolution (1024 × 768) at a refresh rate of 60 Hz. Passages were displayed in Courier New font with 4.76 characters subtending 1° of visual angle. Eye-movements were recorded via an SR Research Eyelink 1000 long-range MRI eyetracker at a sampling rate of 1000 Hz. Viewing was binocular, but only the right eye was tracked.

Procedure

Eyetracking and fMRI data were collected in two 14-min functional runs. Each functional run included 11 paragraph-reading trials, along with 33 filler trials (picture viewing and simple oculomotor scanning, not discussed here). One run contained the paragraphs from *The Emperor's New Clothes*, and the other contained paragraphs from a practice reading assessment test. The order of these runs was counterbalanced across participants. Within both runs, paragraphs were presented in a constant order to maintain story coherence, but paragraph-reading trials were intermixed with filler trials. Each paragraph was presented for 12 s. Participants were instructed to read silently for comprehension as they would normally when reading a novel. An ITI of 6 s was inserted between each trial. Block-level analyses of a subset of these data showed expected activation across the language network (Choi et al., 2014).

Table 1
Means (and standard deviations) for syntactic and lexical measures between the high and the low surprisal conditions.

	High surprisal	Low surprisal
Syntactic surprisal	2.11 (.27)	1.30 (.22)
Lexical surprisal	4.49 (3.72)	4.33 (3.23)
Word frequency	4.59 (1.36)	4.59 (1.39)
Word length	4.23 (2.24)	4.13 (2.22)

Eye-movement data acquisition

A thirteen-point calibration procedure was administered in the scanner before each functional run to map eye position to screen coordinates. Successful calibration required average error less than .49° and maximum error less than .99°. A fixation cross was presented on the screen before each trial, with the first word in the text appearing at that location. Eye movements were recorded throughout each functional run. Stimulus presentation and timing, and eyetracking data collection, were controlled using Experiment Builder software.

fMRI data acquisition

MR data were collected on a Siemens Medical Systems 3T Trio. A 3D T1-weighted "MPRAGE" RF-spoiled rapid flash scan in the sagittal plane, and a T2/PD-weighted multi-slice axial 2D dual Fast Turbo spin-echo scan in the axial plane were used. The multi-echo whole brain T1 scans had 1 mm isotropic voxel size and sufficient field of view to cover from the top of the head to the neck with the following protocol parameters: TR = 2530 ms, TE1 = 1.74 ms, TE2 = 3.6 ms, TE3 = 5.46 ms, TE4 = 7.32 ms, flip angle = 7°. All functional runs were acquired using gradient echo, echo-planar images with the following protocol parameters: TR = 1850 ms, TE = 30 ms, flip angle = 75°. Volumes consisted of 34 three mm slices with transversal orientation. Each volume covered the whole brain with FOV = 208 mm and 64 × 64 matrix, resulting in a voxel size of $3.3 \times 3.3 \times 3\text{mm}^3$.

Eye movement and fMRI co-registration

The fMRI and eyetracking data were synchronized so that fixation onset from the eyetracker could be aligned with the fMRI data. This was accomplished by aligning the onset of the trial run with the onset of the functional scan. Times of experiment onset, block onsets, and fixation onsets were saved in the eye-movement record by Experiment Builder. In addition, both scanner and eyetracker time were recorded via a dedicated TCP/IP port to a separate data logger. This made it possible to co-register eye movement and fMRI events in an event-related analysis.

fMRI analysis

The AFNI software package (Cox, 1996) was used for image analysis. Within-subject analysis involved slice timing correction, spatial co-registration (Cox and Jesmanowicz, 1999) and registration of functional images to the anatomy (Saad et al., 2009). Voxel-wise multiple linear regression was performed with the program 3dREMLfit, using reference functions representing each condition convolved with a standard hemodynamic response function. Reference functions representing the six motion parameters were included as covariates of no interest. In addition, the signal extracted from cerebro-spinal fluid and white matter was included as noise covariates of no interest.

The individual statistical maps and the anatomical scans were projected into standard stereotaxic space (Talairach and Tournoux, 1988) and smoothed with a Gaussian filter of 5 mm FWHM. In a random effects analysis, group maps were created by comparing activations against a constant value of 0. The group maps were thresholded at voxelwise $p < 0.01$ and corrected for multiple comparisons by removing clusters smaller than 658 μl to achieve a mapwise corrected $p < 0.05$. Using the 3dClustSim program with 10,000 iterations, the cluster threshold was determined through Monte Carlo simulations that estimate the chance probability of spatially contiguous voxels exceeding the voxelwise p threshold, i.e., of false positive noise clusters. The smoothness of the data was estimated with the AFNI program 3dFWHMx using regression residuals as input. The analysis was restricted to a mask that excluded areas outside the brain, as well as deep white matter areas and the ventricles.

Table 2
Summary eye movement data. Mean durations (and standard deviations) for each condition averaged over subjects for typical fixation duration measures.

	High surprisal	Low surprisal
First fixation duration (ms)	218 (27)	219 (28)
Single fixation duration (ms)	223 (29)	224 (28)
Gaze duration (ms)	249 (29)	251 (28)
Total fixation time (ms)	278 (33)	280 (37)

Eye movement analysis

The eye-movement data were analyzed off-line using DataViewer (SR Research Ltd., version 1.11.1) to identify fixations and saccades. Saccade detection used three thresholds: the eye had to move at least .1°, with a velocity equal to or greater than 30°/s and with an acceleration of at least $8000°/s^2$. Fixation was defined as a period of time between saccades that was not part of a blink. To be included in the analysis, a fixation could not have a blink immediately before or after, had to fall within a word region, had to have a duration between 50 and 1500 ms, and could not follow a return sweep from the end of one line to the beginning of the next. This resulted in the inclusion of 19,260 fixations across subjects. Basic eye movement measures were examined to ensure that participants were moving their eyes naturally while reading in the scanner (Table 2). All of the eye movement measures were typical of adult readers (Rayner, 1998). Because differences in fixation duration can produce differences in fMRI activation (Henderson and Choi, 2015; Henderson et al., 2015), we compared these measures across the high- and low-surprisal word groups. Importantly, eye movements were very similar across those two groups.

Results

High-surprisal vs. low-surprisal comparison

The results of a whole-brain analysis for the high-surprisal versus low-surprisal comparison are displayed in Table 3 and Fig. 1.[1] Of most direct theoretical interest, as predicted we observed greater activation for the high- than low-surprisal condition in left IFG, and specifically in pars opercularis (BA 44). We also observed activation in left ATL, though it did not reach significance with a voxel-wise threshold level of $p < .01$ (alpha < .05 family-wise error (FWE) corrected). Given our a priori predictions for left ATL, we conducted a secondary analysis in which we used a less conservative voxel-wise threshold of $p < .05$ (alpha < .05 FWE corrected). The results for left IFG were similar to those in the initial analysis, but we also observed the predicted left ATL activation (shown in Fig. 2; peak Talairach coordinates: −46, 7, 12). In contrast to the results for left IFG and left ATL, we observed no evidence for activation in left posterior STS or STG in either analysis.

In addition, although not specifically predicted, we observed greater activation in the main analysis for the high- than low-syntactic surprisal condition bilaterally in the insula, fusiform gyrus, right hemisphere IFG and lingual gyrus, and in two subcortical areas, putamen (bilaterally) and right ventral diencephalon. Only one region, right middle frontal gyrus, showed greater activation in the low-surprisal than high-surprisal condition.

Table 3
Talairach coordinates, volume of the cluster (μl), maximum z-score, and the label of anatomical structure for the high-surprisal vs. low-surprisal analysis, L = left hemisphere, R = right hemisphere.

Volume	Max	x	y	Z	Anatomical structure
High surprisal > low surprisal					
2322	3.711	31	−1	5	R putamen
	3.553	25	13	8	R putamen
	3.178	37	−16	2	R insula, R inferior frontal gyrus (pars opercularis)
2268	4.323	−31	1	0	L insula, L inferior frontal gyrus (pars opercularis), L putamen
837	3.368	28	−43	−9	R fusiform gyrus, R lingual gyrus
810	3.324	−31	−46	−15	L fusiform gyrus
756	3.758	10	−10	−3	R ventral diencephalon
Low surprisal > high surprisal					
729	−4.174	43	19	38	R middle frontal gyrus

Discussion

In this study, we investigated the cortical implementation of syntactic computations during human language comprehension. Subjects read stories presented in paragraph form. Using co-registered eyetracking and fMRI, we measured neural activity as a function of syntactic difficulty as each word was fixated. We operationalized word-by-word syntactic difficulty as syntactic surprisal, which quantifies the likelihood or predictability of a word's syntactic category given its preceding context. Based on prior research, we focused on three cortical regions that have previously been identified as candidates for syntactic computation: left IFG, left ATL, and left posterior STS/STG. We observed greater activation in the left IFG for words with greater syntactic surprisal, consistent with predictions. We also observed a strong trend for greater activation in the left ATL with greater surprisal, also consistent with a priori predictions. In comparison, no evidence for differential activation across surprisal values was observed for left posterior STS/STG. These results are consistent with the hypothesis that processes specifically associated with syntactic difficulty are supported by the left IFG and left ATL. The left IFG (Broca's area) has historically been thought of as the region most closely associated with operations related to syntax. Our results based on syntactic surprisal align with this interpretation. However, we cannot rule out the possibility that other processes beyond syntactic computations that are associated with syntactic surprisal might be the source of the activation in the left LIG. For example, activation in this region may reflect processes associated with differences in syntactic–semantic conflict (Thothathiri et al., 2012), which could correlate with syntactic surprisal.

In the present study, the anterior frontal activation observed in the left IFG extended to increased insula activation for high- versus low-surprisal conditions. Although the insula may not be a traditional focus of theory concerning syntactic computation, its activation is often reported in studies of syntactic processing. For example, a recent Activation Likelihood Estimation (ALE) meta-analysis conducted over 54 empirical studies found that the largest cluster for syntactic processing was the left IFG, with two peaks, one in the pars opercularis and the other in the insula (Rodd et al., 2015). Our results are very much in line with the results of this ALE meta-analysis.

There has also been evidence implicating the left ATL in syntactic processing (Brennan et al., 2010; Dronkers et al., 2004; Friederici et al., 2000a, 2000b, 2003; Humphries et al., 2005, 2006; Mazoyer et al., 1993; Noppeney and Price, 2004; Rogalsky and Hickok, 2009; Vandenberghe et al., 2002). For example, Brennan et al. (2010) investigated the effects of syntactic node count for each word of *Alice in Wonder Land* in a passive listening study and observed left ATL but not left IFG activation. They proposed that the left ATL rather than the left IFG supports syntactic structure building in natural language comprehension. In contrast, in the present study using a similar logic, we observed the clearest relationship between syntactic difficulty and activation in the left IFG, including pars opercularis (BA44) and extending to the insula, with less clear evidence for activation in the left ATL. Given that our study used natural reading without a secondary task, these results suggest that the left IFG plays a central role in syntactic processing during natural language comprehension. At the same time, although the results for the left ATL were not as clear-cut, a liberal voxel-wise alpha

---

[1] An analysis of activation tied to fixation onset, which reflects the average activation across conditions, showed strong activation across the traditional language network.

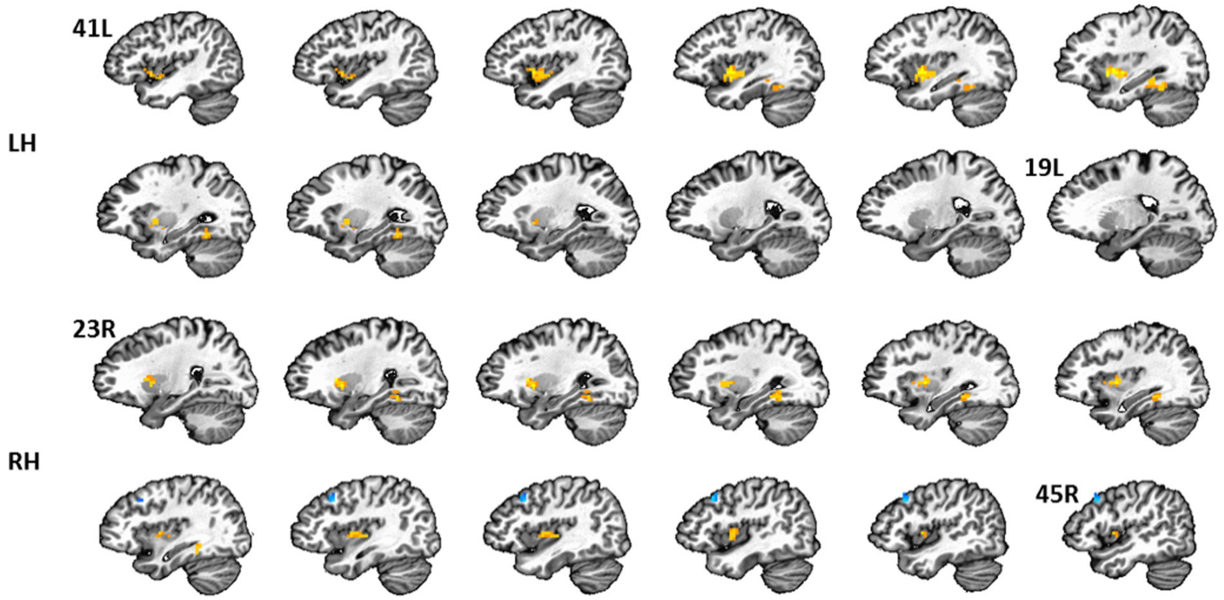


Fig. 1. Areas activated in the contrast of high- versus low-surprisal conditions (whole-brain analysis). Warm regions represent more activation in the high-surprisal condition whereas cool regions reflect more activation in the low-surprisal condition (whole-brain analysis, group maps thresholded at voxelwise $p < 0.01$ and corrected for multiple comparisons by removing clusters with below-threshold size to achieve a mapwise error corrected $p < 0.05$). LH = left hemisphere; RH = right hemisphere. Numbers represent the Talairach x-coordinate of first and last slices.

level along with a standard family-wise error correction for cluster size did show left ATL activation. Given the a priori prediction for activation related to syntax in this region based on the previous literature, these data provide some support for the hypothesis that the left ATL is also involved in syntactic processing. Alternatively, it may be that the left ATL is related to other combinatorial processes that closely track syntactic processing (Zhang and Pylkkänen, 2015) or to prediction more generally (Lau et al., in press).

In comparison to the anterior regions, we did not observe any activation related to syntactic surprisal in the left posterior STS/STG. This region has historically been less associated with syntactic processing compared to anterior regions, though fMRI activation related to syntax is sometimes observed (Constable et al., 2004; Cooke et al., 2002; Hasson et al., 2006; Tyler et al., 2011). Constable et al. (2004) and Cooke et al. (2002) both used an object- versus subject-relative clause comparison in their studies, and Hasson et al. (2006) compared subordinate-clause sentences (many of which included relative clauses) to controls. It may be that these comparisons produce a larger syntactic effect. Alternatively, the need to resolve complex clausal structures also requires additional syntactic and semantic reanalysis and integration processes. In line with this interpretation, Grodzinsky and Friederici (2006) concluded in their review of the literature that the left posterior STS/STG plays a role in final syntactic integration following syntactic computations initially supported by anterior regions. Of course, the lack of posterior STS/STG activation in the current study represents a failure to reject the null hypothesis, so caution is warranted: We cannot rule out the possibility that the lack of left posterior STS/STG activation was due to Type II error. However, we note that the study design was of sufficient statistical power to produce syntactic effects in other cortical regions, suggesting that the experimental technique is generally capable of detecting influences of complexity.

Two recently published fMRI studies of surprisal in language processing during story listening reported related results (Hale et al., 2015; Willems et al., in press; see also Bachrach, 2008, for a related project). Although the general strategy of using surprisal and fMRI in these studies was similar to that taken in the present study, they differed in a number of critical ways, including the input modality (speech versus written text) and the type of surprisal investigated. In the present study, we focused specifically on syntactic surprisal for part of speech derived from a lexicalized probabilistic context free grammar. Willems et al.'s measure of surprisal was based on word co-occurrence frequencies using a trigram model. The trigram metric does not take syntax directly into account, and instead focuses on general lexical constraint. In addition, lexical co-occurrence is typically highly correlated with lexical frequency, such that higher co-occurrence values are seen for words with higher base frequency counts. Lexical frequency itself generates strong fMRI activity in cortical language areas (Graves et al., 2010; Hauk et al., 2008; Price, 2012). Importantly for isolating effects of syntactic complexity, the syntactic surprisal conditions in the present study controlled for lexical frequency, word length, and word predictability, three of the most important variables in predicting natural reading behavior (Clifton et al., 2016). The present syntactic surprisal effects

Fig. 2. Left hemisphere anterior temporal lobe activation in the contrast of high- versus low-surprisal conditions (whole-brain analysis, group maps thresholded at voxelwise $p < 0.05$ and corrected for multiple comparisons by removing clusters with below-threshold size to achieve a mapwise error corrected $p < 0.05$). Numbers represent the Talairach x-coordinate of the slice.

therefore cannot be attributed to these other variables. It will be important in future studies to directly compare the effects of different types of surprisal and complexity metrics computed over a common text corpus.

More consistent with the present study's focus on syntactic difficulty, Hale et al. (2015) focused on linguistic prediction. They presented *Alice in Wonderland* via speech and examined linguistic predictors across representational grain size. They observed left IFG activation for surprisal based on lexical (bi-gram) co-occurrence, and bilateral anterior ATL activation for syntactic difficulty derived from phrase structure grammars and from syntactic node counts generated from Minimalist Grammar.

In addition to observing activation related to syntactic surprisal in the predicted language regions, we also observed greater activation for the high versus low syntactic surprisal conditions in a number of other regions: bilateral insula, fusiform gyrus, and putamen; and right hemisphere IFG (bilateral IFG including the predicted left IFG activation already discussed), lingual gyrus, and ventral diencephalon. First, the insula has been associated with a variety of levels of language processing including production (e.g., Dronkers, 1996; McCarthy et al., 1993; Price, 2012), auditory processing (Bamiou et al., 2003), and syntactic processing (Moro et al., 2001). For example, Moro et al. (2001) reported that the left insula was more activated when participants covertly read syntactically illegal sentences relative to when they read phonotactically illegal sentences, consistent with greater activation for higher syntactic surprisal in the present study. Second, fusiform and neighboring lingual gyrus are involved in visual letter and word form processing and semantic processing during language comprehension (Price, 2012). Although there is no direct evidence for a relationship between syntactic processing and fusiform gyrus, our result may suggest that words with greater syntactic surprisal elicit additional orthographic encoding or greater semantic processing at the lexical level, either as a check on lexical encoding or due to additional integration processes associated with syntactic reanalysis. Third, the putamen has typically been found to be related to motor skills including language production (Price, 2012). In addition, and consistent with our results, some studies have reported that the putamen and adjacent areas are associated with syntactic and semantic processes in language comprehension (Lieberman, 2001; Pickett et al., 1998). Fourth, greater activation in the high than in the low surprisal conditions in bilateral diencephalon including thalamus is consistent with Wahl et al. (2008), who examined the role of thalamic regions in syntactic processing using subcortical Deep Brain Stimulation. They found that thalamic areas systematically responded to syntactic and semantic errors during auditory language comprehension. Fifth, the observed right hemisphere IFG activation (along with the left IFG activation) is consistent with the bilateral IFG activation sometimes observed in tasks that reflect linguistic expectations (Bonhage et al., 2015; Wlotko and Federmeier, 2007).

Finally, we observed greater activation of right middle frontal gyrus (MFG) for the low relative to high syntactic surprisal conditions. Right MFG is typically associated with working memory, executive function, and inhibitory control. The finding of greater activity in this region in the low compared to high syntactic surprisal condition was not expected. We speculate that the effect may be related to attentional processes associated with right MFG function. During reading, attention is typically programmed to shift to the next word as analysis succeeds on the current word (Henderson and Ferreira, 1990; Reichle et al., 2003). However, if syntactic analysis is more difficult, as is the case in the higher syntactic surprisal condition, this controlled shift of attention may be delayed, leading to less activity tied to the current fixation in the high than low surprisal condition.

It is interesting to note that the effects of syntactic difficulty seen in the fMRI data were not apparent in the eyetracking data. An obvious question is why we did not observe surprisal effects in the eyetracking measures given past reports of surprisal effects on reading time (Boston et al., 2008; Demberg and Keller, 2008; Demberg et al., 2013; Levy, 2008). It is important to note, however, that these earlier studies investigated either overall surprisal (i.e., a combination of n-gram and lexicalized syntactic surprisal) or un-lexicalized part-of-speech (POS) surprisal, both of which are different from the surprisal measure we investigated. In contrast, it is less clear exactly how reading times relate to the lexicalized syntactic surprisal metric we used here. Roark et al. (2009) reported a significant effect of this type of surprisal on self-paced reading times, but only for content words once function words had been removed from their analysis. Importantly, however, Roark et al. used self-paced reading, which is slower and tends to index later and more strategic language processes compared to eyetracking. In addition, the difference in results could be due to differences in the stories used in Roark et al. versus in our study. Roark et al. used materials created specifically to examine syntactic processes (Bachrach, 2008) and included many sentences that tended to be relatively complex and difficult. In comparison, we used passages that were created independently of our study and that were written simply to be read. It is possible that the syntactic surprisal values of our stimuli are more representative of what people would typically encounter in day-to-day linguistic contexts. We hypothesize that the combination of differences in stimulus materials and reading time methods may be at least partly responsible for the difference in reading time results across studies. We note also that the lack of any effect of syntactic surprisal in the two middle surprisal quartiles that we included in the fMRI analysis was borne out over the entire range of syntactic surprisal values: In an unreported analysis we found no significant reading time effects in the eyetracking record as a function of syntactic surprisal over the full data set. At the same time, we did observe clear overall (n-gram plus syntactic) surprisal effects on reading time, consistent with previous literature.

We note that although syntactic surprisal captures an important aspect of syntactic difficulty, it is only one method for assessing syntactic processing. For example, Hale et al. (2015) and Brennan et al. (2010) operationalized syntactic complexity as a count of the number of syntactic nodes needed to integrate each word into the phrase structure, relating difficulty to the density of syntactic nodes associated with a word (Frazier, 1985; Hawkins, 1994). Another approach might be to implement a metric based on the dependency-locality theory, which relates processing difficulty to the amount of integration required and the distance over which that integration must take place (Gibson, 1998, 2000). The fixation-related fMRI approach for studying natural reading used here provides a new method for investigating and comparing the effects of these alternative measures of syntactic complexity.

In conclusion, from a theoretical perspective, our results are important in showing that the psycholinguistic construct of syntactic surprisal produces a clear neurocognitive effect during reading. Indeed, syntactic surprisal modulates activity in two areas of the brain typically associated with syntactic processing. More generally, syntactic surprisal produces activation in cortical regions that have been associated with prediction in language. These results provide evidence for the neurocognitive reality of the theoretical construct of syntactic surprisal. From a methodological perspective, the clear fMRI effects that we observed in the absence of eye movement effects suggest that FIRE fMRI can provide a unique source of evidence for testing theoretical questions in language processing during reading, supplementing the behavioral methods often used in psycholinguistics. These results in turn suggest that FIRE fMRI can provide an important source of evidence for testing theories of language representation and computation "in the wild" without the need to examine highly difficult (and often unnatural) syntactic forms.

Acknowledgments

We thank Vera Demberg-Winterfors for her insights concerning surprisal parsers, Brian Roark for guidance with his parser, and Rutvik Desai for his fundamental contributions to developing FIRE fMRI. We also thank Tamara Swaab, Ellen Lau, John Hale, and the Language, Cognition and Brain Sciences group at UC Davis for their helpful feedback on this work. This research was supported by the National Science Foundation (BCS-1151358). The authors declare no competing financial interests.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2016.02.050.

References

Altmann, G.T.M., Kamide, Y., 1999. Incremental interpretation at verbs: restricting the domain of subsequent reference. Cognition 73, 247–264.

Altmann, U., Bohrn, I.C., Lubrich, O., Menninghaus, W., Jacobs, A.M., 2014. Fact vs fiction—how paratextual information shapes our reading processes. Soc. Cogn. Affect. Neurosci. 9, 22–29.

Bachrach, A., 2008. Imaging neural correlates of syntactic complexity in a naturalistic context Ph.D. thesis Massachusetts Institute of Technology, Cambridge, Massachusetts.

Bamiou, D.-E., Musiek, F.E., Luxon, L.M., 2003. The insula (island of reil) and its role in auditory processing: literature review. Brain Res. Rev. 42, 143–154.

Bar, M., 2009. Predictions: a universal principle in the operation of the human brain. Philos. Trans. R. Soc. Biol. Sci. 364, 1181–1182.

Bartels, A., Zeki, S., 2004. Functional brain mapping during free viewing of natural scenes. Hum. Brain Mapp. 21, 75–85.

Ben-Shachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., Grodzinsky, Y., 2003. The neural reality of syntactic transformations: evidence from fMRI. Psychol. Sci. 14, 433–440.

Ben-Shachar, M., Palti, D., Grodzinsky, Y., 2004. Neural correlates of syntactic movement: converging evidence from two fMRI experiments. NeuroImage 21 (4), 1320–1336.

Bonhage, C.E., Mueller, J.L., Friederici, A.D., Fiebach, C.J., 2015. Combined eye tracking and fMRI reveals neural basis of linguistic predictions during sentence comprehension. Cortex 68, 33–47.

Bornkessel-Schlesewsky, I., Schlesewsky, M., von Cramon, D.Y., 2009. Word order and Broca's region: evidence for a supra-syntactic perspective. Brain Lang. 111, 125–139.

Boston, M.F., Hale, J.T., Patil, U., Kliegl, R., Vasishth, S., 2008. Parsing costs as predictors of reading difficulty: an evaluation using the Potsdam sentence corpus. J. Eye Mov. Res. 2, 1–12.

Brennan, J., Nir, Y., Hasson, U., Malach, R., Heeger, D., Pylkkänen, L., 2010. Syntactic structure building in the anterior temporal lobe during natural story listening. Brain Lang. 120, 163–173.

Brysbaert, M., New, B., 2009. Moving beyond Kučera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. Behav. Res. Methods 41, 977–990.

Caplan, D., Alpert, N., Waters, G., Olivieri, A., 2000. Activation of Broca's area by syntactic processing under conditions of concurrent articulation. Hum. Brain Mapp. 9, 65–71.

Caplan, D., Chen, E., Waters, G., 2008. Task-dependent and task-independent neurovascular responses to syntactic processing. Cortex 44 (3), 257–275.

Carramazza, A., Zurif, E., 1976. Dissociations of algorithmic and heuristic processes in sentence comprehension: evidence from aphasia. Brain Lang. 3, 572–582.

Choi, W., Henderson, J.M., 2015. Neural correlates of active vision: an fMRI comparison of natural reading and scene viewing. Neuropsychologia 75, 109–115. http://dx.doi.org/10.1016/j.neuropsychologia.2015.05.027.

Choi, W., Desai, R.H., Henderson, J.M., 2014. The neural substrates of natural reading: a comparison of normal and nonword text using eyetracking and fMRI. Front. Hum. Neurosci. 8, 1024. http://dx.doi.org/10.3389/fnhum.2014.01024.

Clark, A., 2013. Whatever next? PRedictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. 36, 181–204.

Clifton, C.E., Ferreira, F., Henderson, J.M., Inhoff, A.W., Liversedge, S., Reichle, E.D., Schotter, E.R., 2016. Eye movements in reading and information processing: Keith Rayner's 40 year legacy. J. Mem. Lang. 86, 1–19.

Constable, R.T., Pugh, K.R., Berroya, E., Mencl, W.E., Westerveld, M., Ni, W., et al., 2004. Sentence complexity and input modality effects in sentence comprehension: an fMRI study. NeuroImage 22 (1), 11–21.

Cooke, A., Zurif, E.B., DeVita, C., Alsop, D., Koenig, P., Detre, J., et al., 2002. Neural basis for sentence comprehension: grammatical and short-term memory components. Hum. Brain Mapp. 15 (2), 80–94.

Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput. Biomed. Res. 29, 162–173.

Cox, R.W., Jesmanowicz, A., 1999. Real-time 3D image registration of functional MRI. Magn. Reson. Imaging 42, 1014–1018.

Dapretto, M., Bookheimer, S.Y., 1999. Form and content: dissociating syntax and semantics in sentence comprehension. Neuron 24 (2), 427–432.

Demberg, V., Keller, F., 2008. Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. Cognition 109, 193–210.

Demberg, V., Keller, F., Koller, A., 2013. Incremental, predictive parsing with psycholinguistically motivated tree-adjoining grammar. Comput. Linguist. 39, 1025–1066.

Den Ouden, H.E.M., Kok, P., de Lange, F.P., 2012. How prediction errors shape perception, attention, and motivation. Front. Psychol. 3 (Article 548).

Dronkers, N.F., Wilkins, D.P., Van sVralin Jr., R.D., Redfern, B.B., Jaeger, J.J., 2004. Lesion analysis of the brain areas involved in language comprehension. Cognition 92 (1–2), 145–177.

Dronkers, N.F., 1996. A new brain region for coordinating speech articulation. Nature 384, 159–161.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., Sakai, K.L., 2000. A syntactic specialization for Broca's Area. Proc. Natl. Acad. Sci. 97 (11), 6150–6154.

Fedorenko, E., Nieto-Castañon, A., Kanwisher, N., 2012. Syntactic processing in the human brain: what we know, what we don't know, and a suggestion for how to proceed. Brain Lang. 120 (2), 187–207. http://dx.doi.org/10.1016/j.bandl.2011.01.001.

Frank, S.L., Otten, L.J., Galli, G., Vigliocco, G., 2015. The ERP response to the amount of information conveyed by words in sentences. Brain Lang. 140, 1–11.

Frazier, L., 1985. Syntactic Complexity. In: Dowty, D., Karttunen, L., Zwicky, A.M. (Eds.), Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives. Cambridge University Press, Cambridge, pp. 129–187.

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

Friederici, A.D., Fiebach, C.J., Schlesewsky, M., Bornkessel, I.D., von Cramon, D.Y., 2006. Processing linguistic complexity and grammaticality in the left frontal cortex. Cereb. Cortex 16 (12), 1709–1717.

Friederici, A.D., Meyer, M., von Cramon, D.Y., 2000a. Auditory languge comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang. 74, 289–300.

Friederici, A.D., Ruschenmeyer, S.-A., Hahne, A., Fiebach, C.J., 2003. The role of left inferior frontal and Superior temporal cortex in sentence comprehension: localizing syntactic and semantic processes. Cereb. Cortex 13, 170–177.

Friederici, A.D., Wang, Y., Herrmann, C.S., Maess, B., Oertel, U., 2000b. Localization of early syntactic processes in frontal and temporal cortical areas: a magnetoencephalographic study. Hum. Brain Mapp. 11, 1–11.

Friston, K.J., 2010. The free-energy principle: A unified brain theory? Nat. Rev. Neurosci. 11, 127–138.

Gibson, E., 1998. Linguistic complexity: locality of syntactic dependencies. Cognition 68, 1–76.

Gibson, E., 2000. The dependency locality theory: a distance-based theory of linguistic complexity. In: Miyashita, Y., Marantz, A., O'Neil, W. (Eds.), Image, Language, Brain. MIT Press, Cambridge, pp. 95–126.

Graves, W.W., Desai, R., Humphries, C., Seidenberg, M.S., Binder, J.R., 2010. Neural systems for reading aloud: a multiparametric approach. Cereb. Cortex 20 (8), 1799–1815.

Grodzinsky, Y., 2001. The neurology of syntax: language use without Broca's area. Behav. Brain Sci. 23 (01), 1–21.

Grodzinsky, Y., Friederici, A.D., 2006. Neuroimaging of syntax and syntactic processing. Curr. Opin. Neurobiol. 16 (2), 240–246.

Hale, J., 2001. A probabilistic Earley parser as a psycholinguistic model. Proc. NAACL 2. Association for Computational Linguistics, pp. 159–166. http://dx.doi.org/10.3115/1073336.1073357.

Hale, J.T., Lutz, D.E., Luh, W.-M., Brennan, J.R., 2015. Modeling fMRI time courses with linguistic structure at various grain sizes. Proc. CMCL, pp. 89–97.

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303, 1634–1640.

Hasson, U., Nusbaum, H.C., Small, S.L., 2006. Repetition suppression for spoken sentences and the effect of task demands. J. Cogn. Neurosci. 18 (12), 2013–2029.

Hauk, O., Davis, M.H., Pulvermüller, F., 2008. Modulation of brain activity by multiple lexical and word form variables in visual word recognition: a parametric fMRI study. NeuroImage 42 (3), 1185–1195.

Hawkins, J.A., 1994. A Performance Theory of Order and Constituency. Cambridge University Press.

Henderson, J.M., Choi, W., 2015. Neural correlates of fixation duration during real-world scene viewing: evidence from fixation-related (FIRE) fMRI. J. Cogn. Neurosci. 27 (6), 1137–1145.

Henderson, J.M., Ferreira, F., 1990. Effects of foveal processing difficulty on the perceptual span in reading: implications for attention and eye movement control. J. Exp. Psychol. Learn. Mem. Cogn. 16, 417–429.

Henderson, J.M., Choi, W., Luke, S.G., Desai, R.H., 2015. Neural correlates of fixation duration in natural reading: evidence from fixation-related fMRI. NeuroImage 119, 390–397.

Humphries, C., Binder, J.R., Medler, D.A., Liebenthal, E., 2006. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. J. Cogn. Neurosci. 18, 665–679.

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

Just, M.A., Carpenter, P.A., Keller, T.A., Eddy, W.F., Thulborn, K.R., 1996. Brain activation modulated by sentence comprehension. Science 274 (5284), 114.

Kaan, E., Swaab, T.Y., 2002. The brain circuitry of syntactic comprehension. Trends Cogn. Sci. 6, 350–356.

Kamide, Y., Altmann, G.T.M., Haywood, S.L., 2003. The time-course of prediction in incremental sentence processing: evidence from anticipatory eye movements. J. Mem. Lang. 49, 133–156.

Kuperberg, G.R., Jaeger, T.F., 2016. What do we mean by prediction in language comprehension? Lang. Cogn. Neurosci. 31 (1), 32–59.

Kutas, M., DeLong, K.A., Smith, N.J., 2011. A look around at what lies ahead: prediction and predictability in language processing. In: Bar, M. (Ed.), Predictions in the brain: using our past to generate a future. Oxford University Press, pp. 190–207.

Lau, E.F., Weber, K., Gramfort, A., Hamalainen, M.S., Kuperberg, G.R., 2016. Spatiotemporal signatures of lexical-semantic prediction. Cereb. Cortex (in press).

Levy, R., 2008. Expectation-based syntactic comprehension. Cognition 106, 1126–1177.

Lieberman, P., 2001. Human language and our reptilian brain: the subcortical bases of speech, syntax, and thought. Perspect. Biol. Med. 44, 32–51.

Lupyan, G., Clark, A., 2015. Words and the world: predictive coding and the language-perception-cognition interface. Curr. Dir. Psychol. Sci. 24, 279–284.

Mani, N., Huettig, F., 2012. Prediction during language processing is a piece of cake—but only for skilled producers. J. Exp. Psychol. Hum. Percept. Perform. 38, 843–847.

Marcus, M.P., Marcinkiewicz, M.A., Santorini, B., 1993. Building a large annotated corpus of English: the Penn treebank. Comput. Linguist. 19, 313–330.

Marsman, J.B.C., Renken, R., Velichkovsky, B.M., Hooymans, J.M.M., Cornelissen, F.W., 2012. Fixation based event-related fMRI analysis: using eye fixations as events in functional magnetic resonance imaging to reveal cortical processing during the free exploration of visual images. Hum. Brain Mapp. 33, 307–318.

Mazoyer, B.M., Tzourio, N., Frak, V., Syrota, A., Murayama, N., Levrier, O., Salamon, G., Dehaene, S., Cohen, L., Mehler, J., 1993. The cortical representation of speech. J. Cogn. Neurosci. 5, 467–479.

McCarthy, G., Blamire, A.M., Rothman, D.L., Gruetter, R., Shulman, R.G., 1993. Echo-planar magnetic resonance imaging studies of frontal cortex activation during word generation in humans. Proc. Natl. Acad. Sci. U. S. A. 90, 4952–4956.

Misyak, J.B., Christiansen, M.H., Tomblin, J.B., 2010. Sequential expectations: the role of prediction-based learning in language. Top. Cogn. Sci. 2, 138–153.

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S.F., Fazio, F., 2001. Syntax and the brain: disentangling grammar by selective anomalies. NeuroImage 13, 110–118.

Nishimoto, S., Vu, A.T., Naselaris, T., Benjamini, Y., Yu, B., Gallant, J.L., 2011. Reconstructing visual experiences from brain activity evoked by natural movies. Curr. Biol. 21, 1641–1646.

Noppeney, U., Price, C.J., 2004. An fMRI study of syntactic adaptation. J. Cogn. Neurosci. 16 (4), 702–713.

Pickering, M.J., Garrod, S., 2004. Toward a mechanistic psychology of dialogue. Behav. Brain Sci. 27, 169–190.

Pickering, M.J., Garrod, S., 2013. An integrated theory of language production and comprehension. Behav. Brain Sci. 36, 329–392.

Pickett, E.R., Kuniholm, E., Protopapas, A., Friedman, J., Lieberman, P., 1998. Selective speech motor, syntax, and cognitive deficits associated with bilateral damage to the putamen and the head of the caudate nucleus: a case study. Neuropsychologia 36, 173–188.

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 (2), 816–847.

Rao, R.P.N., Ballard, D.H., 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2, 79–87.

Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. Psychol. Bull. 124, 372–422.

Reichle, E.D., Rayner, K., Pollatsek, A., 2003. The EZ reader model of eye-movement control in reading: comparisons to other models. Behav. Brain Sci. 26 (04), 445–476.

Richlan, F., Gagl, B., Hawelka, S., Braun, M., Schurz, M., Kronbichler, M., et al., 2014. Fixation-related fMRI analysis in the domain of reading research: using self-paced eye movements as markers for hemodynamic brain responses during visual letter string processing. Cereb. Cortex 24, 2647–2656.

Roark, B., 2001. Probabilistic top-down parsing and language modeling. Comput. Linguist. 27, 249–276.

Roark, B., Bachrach, A., Cardenas, C., Pallier, C., 2009. Deriving lexical and syntactic expectation-based measures for psycholinguistic modeling via incremental top-down parsing. Proceedings of the 2009 Conference for Psycholinguistic Modeling via Incremental Top-Down Parsing, pp. 324–333.

Rodd, J.M., Vitello, S., Woollams, A.M., Adank, P., 2015. Localising semantic and syntactic processing in spoken and written language comprehension: an activation likelihood estimation meta-analysis. NeuroImage 141, 89–102.

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

Saad, Z.S., Glen, D.R., Chen, G., Beauchamp, M.S., Desai, R., Cox, R.W., 2009. A new method for improving functional-to-structural MRI alignment using local Pearson correlation. NeuroImage 44, 839–848.

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

Smith, N.J., Levy, R., 2013. The effect of word predictability on reading time is logarithmic. Cognition 128, 302–319.

Stansbury, D.E., Naselaris, T., Gallant, J.L., 2013. Natural scene statistics account for the representation of scene categories in human visual cortex. Neuron 79, 1025–1034.

Staub, A., 2015. The effect of lexical predictability on eye movements in reading: critical review and theoretical interpretation. Lang Ling Compass 9, 311–327.

Stromswold, K., Caplan, D., Alpert, N., Rauch, S., 1996. Localization of syntactic comprehension by positron emission tomography. Brain Lang. 52, 452–473.

Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme Medical, New York.

Thothathiri, M., Kim, A., Trueswell, J.C., Thompson-Schill, S.L., 2012. Parametric effects of syntactic–semantic conflict in Broca's area during sentence processing. Brain Lang. 120 (3), 259–264.

Tyler, L.K., Marslen-Wilson, W.D., Randall, B., Wright, P., Devereux, B.J., Zhuang, J., ... Stamatakis, E.A., 2011. Left inferior frontal cortex and syntax: function, structure and behaviour in patients with left hemisphere damage. Brain J. Neurol. 134 (Pt 2), 415–431 http://doi.org/10.1093/brain/awq369.

Van Berkum, J.J.A., Brown, C.M., Zwisterlood, P., Kooijman, V., Hagoort, P., 2005. Anticipating upcoming words in discourse: evidence from ERPs and reading times. J. Exp. Psychol. Learn. Mem. Cogn. 31, 443–467.

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

Wahl, M., Marzinzik, F., Friederici, A.D., Hahne, A., Kupsch, A., Schneider, G.-H., et al., 2008. The human thalamus processes syntactic and semantic language violations. Neuron 59, 695–707.

Wicha, N.Y.Y., Moreno, E.M., Kutas, M., 2003. Expecting gender: an event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. Cortex 39, 483–508.

Willems, R.M., Frank, S.L., Nijhof, A.D., Hagoort, P., van den Bosch, A., 2016. Prediction during natural language comprehension. Cereb. Cortex (in press).

Wilson, S.M., Molnar-Szakacs, I., Iacoboni, M., 2008. Beyond superior temporal cortex: intersubject correlations in narrative speech comprehension. Cereb. Cortex 18 (1), 230–242.

Wlotko, E.W., Federmeier, K.D., 2007. Finding the right word: hemispheric asymmetries in the use of sentence context information. Neuropsychologia 45 (13), 3001–3014.

Zhang, L., Pylkkänen, L., 2015. The interplay of composition and concept specificity in the left anterior temporal lobe: An MEG study. NeuroImage 111, 228–240.