Prediction during natural language comprehension

Roel M. Willems[1,2], Stefan L. Frank[3], Annabel D. Nijhof[1], Peter Hagoort[1,2], Antal van den Bosch[1,3]

1. Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands

2. Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

3. Centre for Language Studies, Radboud University Nijmegen, The Netherlands

For correspondence:

Roel Willems

Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen

P.O. Box 9101

6500 HB Nijmegen

The Netherlands

Tel.: +31 24 361 4668

roel.willems@donders.ru.nl

Abstract

The notion of *prediction* is increasingly studied in cognitive neuroscience. We investigated the neural basis of two distinct aspects of word prediction, derived from information theory, during story comprehension. We assessed the effect of *entropy* of next-word probabilities as well as *surprisal*. A computational model determined entropy and surprisal for each word in three literary stories. Twenty-four healthy participants listened to the same three stories while their brain activation was measured using fMRI. Reversed speech fragments were presented as a control condition.

Brain areas sensitive to entropy were left ventral premotor cortex, left middle frontal gyrus, right inferior frontal gyrus, left inferior parietal lobule, and left supplementary motor area. Areas sensitive to surprisal were left inferior temporal sulcus ('visual word form area'), bilateral superior temporal gyrus, right amygdala, bilateral anterior temporal poles, and right inferior frontal sulcus. We conclude that prediction during language comprehension can occur at several levels of processing, including at the level of word form. Our study exemplifies the power of combining computational linguistics with cognitive neuroscience, and additionally underlines the feasibility of studying continuous spoken language materials with fMRI.

Keywords: Prediction; Word surprisal; Entropy; Language; fMRI

Introduction

It has become increasingly clear that the brain should be seen as a proactive organ, actively predicting what will happen next, instead of being a passive input-chewing device (e.g. Friston 2005; Bar 2009; den Ouden et al. 2012; Clark 2013). Prediction is a powerful mechanism, allowing for the mental speed that smooth cognitive functioning requires. During language comprehension, too, there is evidence for prediction. For instance, comprehenders actively predict an upcoming word, when that word is predictable from the preceding context (Wicha et al. 2004; DeLong et al. 2005; Van Berkum et al. 2005; Federmeier 2007; Dambacher et al. 2009; Laszlo and Federmeier 2009; Dikker et al. 2010; Dikker and Pylkkänen 2013; Lau et al. 2014). Although prediction has not been investigated as much in the language domain as in other domains of cognitive neuroscience, the research that is available indicates that prediction plays a role during language comprehension (see Van Petten and Luka 2012 for overview).

In this study we investigated the effects of prediction on the neural language network. Participants' brain activation was measured using fMRI while they listened to spoken narratives. Prediction was quantified by means of a computational linguistic model that assigned occurrence probabilities to all words that might come next at each point in the narrative. The model then estimated two indices related to word prediction. First, the model estimated the *entropy* of the distribution of next-word probabilities; a measure that quantifies how uncertain the model is about what will come next. Second, the model estimated *surprisal*, which expresses how unexpected the current word is given the previously encountered words.

Although both entropy and surprisal can be taken as measures of word prediction, they quantify different concepts. Entropy is high when many different words may occur next, that is, the *upcoming* word is hard to predict from the text so far. In contrast, surprisal is high when the current word was unexpected, that is, it did not conform with the prediction. In other words, entropy is forward-looking whereas surprisal is backward-looking.

We will now introduce entropy and word surprisal as concepts from information theory more fully, before turning to our neural hypotheses.

Surprisal and entropy

A sentence or text can simply be formalized as a sequence of words: w_1, w_2, … . We assume that the language-comprehension system, after processing the first $t-1$ words (i.e., the sequence $w_1,…,w_{t-1}$), is in a state that implicitly assigns a conditional probability $P(w_t|w_1,…,w_{t-1})$ to each potentially upcoming word w_t. The surprisal associated with observing the word that actually appears at position t is defined as the negative logarithm of its occurrence probability:

$$\text{surprisal}(t) = -\log P(w_t | w_1,\ldots,w_{t-1}).$$

If the observed word's probability equals 1 (i.e., no other word was considered possible given the preceding context), its surprisal equals 0. Conversely, the occurrence of a word that was not among the words considered possible (i.e., has zero probability) corresponds to infinite surprisal. Surprisal can be thought of as the degree to which the actually perceived word w_t deviates from expectation; this interpretation highlights the importance of prediction for word surprisal. Word surprisal is formally identical to self-information, and is sometimes referred to simply as 'surprise'.

The measure of word surprisal has proved to be very powerful, e.g. as an optimization criterion in the decoders of statistical machine translation (Koehn 2010). Also, word surprisal has been found to predict the length of words, with shorter words being used in less surprising situations (Piantadosi et al. 2011; Mahowald et al. 2013). An important issue is whether word surprisal accurately captures cognitive processing during language processing. Hale (2001) and Levy (2008) argue that integrating a word into the current context requires an amount of cognitive processing effort that is proportional to the word's surprisal. If surprisal indeed quantifies language processing effort, it should correlate with experimental measures of comprehension difficulty. Several previous studies in which word surprisal estimates were compared to data from sentence reading experiments confirm that surprisal indeed correlates positively with reading time (Monsalve et al. 2012; Frank and Thompson 2012; Frank 2013; Smith and Levy 2013). Moreover, it was found that the amplitude of the N400 event-related potential (ERP) component correlates with word surprisal values (Frank et al. 2015) . The fact that surprisal correlates with the amplitude of a classical ERP component related to language comprehension (Kutas and Federmeier 2011) is another source of evidence for the hypothesis that surprisal indeed captures aspects of language comprehension.

The second information-theoretic quantity we investigate here, entropy, is also derived from the conditional probabilities of words given the text so far. However, unlike surprisal, it is not a function of the *current* word's probability but of the distribution of probabilities of all possible *upcoming* words. It is defined as:

$$\text{entropy}(t) = -\sum_{w_{t+1} \in W} P(w_{t+1}\, | w_1, \ldots, w_t) \log P(w_{t+1} | w_1, \ldots, w_t),$$

where W denotes the set of all word types.

Note that the definition of *surprisal* at position *t* makes use of the probability of the word w_t, whereas the *entropy* at position *t* depends on the probabilities of potentially upcoming words w_{t+1}. If the context $w_1,\ldots,w_t$ is not very predictive about w_{t+1}, the total probability is distributed over many words, resulting in high entropy. Conversely, if only a small set of words is likely to follow the current context, many words will have (near) zero probability and entropy is low. In the extreme case where a single word is considered to occur with certainty, entropy equals zero.

Only very few studies have looked at behavioural or neural effects of entropy during language comprehension, with mixed results. No correlation has been found between entropy(*t*) and reading time (Frank, 2013) or ERP amplitude (Frank et al., 2015) on w_{t+1} (at least, not after factoring out the effect of surprisal of w_{t+1}). That is, uncertainty about the upcoming word does not appear to affect processing of that word as indexed by reading times and ERPs. However, Roark et al. (2009) found that w_t is read more slowly when entropy(*t*) is higher, suggesting that *entering* a state of high uncertainty slows down processing. The current study, too, investigates the relation between entropy(*t*) and processing of w_t but looks at brain activation rather than behavioural measures.

A remaining question then is: What are the neural areas sensitive to entropy and surprisal during language comprehension? The present study sets out to answer this issue, and specifically looks into the stages of the cortical hierarchy that are influenced by these measures of word prediction.

The current study

In the current study we want to add to the existing literature in three ways. First, we investigate which brain areas are involved in entropy and surprisal of words during comprehension of spoken language stimuli. A remaining issue is at what level of neural processing prediction occurs during language processing. If word processing conforms to the principles of predictive coding, surprisal should be expressed throughout the auditory (or language) hierarchy as predictions at higher level descend to lower levels to incorporate prediction errors (e.g. Friston 2005). In our setting, surprisal reports the prediction error or unpredicted aspects of a stimulus. In predictive coding schemes, the predictability or precision of predictions amplifies prediction errors. This priming or (synaptic) gain control is consistent with modulation of early cortical areas (such as primary visual cortex). Indeed, modulations of early cortical areas by predictability have been found in the domain of visual perception (e.g. Kok et al. 2012), as well as in magneto- or electro-encephalography studies of language comprehension (Dambacher et al. 2009; Dikker et al. 2010; Dikker and Pylkkänen 2013). If predictability leads to pre-activation at the level of word form (as suggested by a predictive coding framework), we predict to see an effect in areas sensitive to word form processing, or other parts of early sensory cortex (Dikker et al. 2010).

Prediction may also influence areas more generally thought to be implicated in integrative processes during language processing. Candidate regions are the left and right inferior frontal gyri (IFG), given that they are known to play an important role in integration during sentence and discourse comprehension (e.g. Robertson et al. 2000; Mason and Just 2004; Ferstl et al. 2008; Hagoort et al. 2009; Menenti et al. 2009). Specifically, Hagoort (2005, 2013) hypothesized that IFG acts as a 'unification space' for language, meaning that it plays a role in preselection as well as integration of upcoming / perceived information. The anterior temporal poles are other candidate regions given their sensitivity to predictability of context (e.g. Lau et al. 2014). Note that these two scenarios (modulation of areas early in the cortical hierarchy and of more 'integrative' areas) are not mutually exclusive.

Second, we aim at separating effects of surprisal and entropy. These two sides of prediction have been shown to have separable neural effects in studies using non-language stimuli (e.g. Strange et al. 2005; Tobia et al. 2012; Ahlheim et al. 2014; Nastase et al. 2014), and here we investigate whether a similar distinction is present in the language domain.

Finally, this study extends previous research in using extended narratives as stimuli. Our participants listened to full spoken narratives presented at a natural speed, without an artificial experimental task. This means we test effects of prediction in more natural settings than is usually done (such as by presenting single sentences). The present study falls within a growing body of research which investigates language processes with more naturalistic stimuli such as narratives (Nijhof and Willems, 2015; Speer et al. 2009; Lerner et al. 2011; Wallentin et al. 2011; Brennan et al. 2012; Kurby and Zacks 2013; Altmann et al. 2014; Hsu et al. 2014; Jacobs 2015).

Methods

Participants

Twenty-four healthy, native speakers of Dutch (8 male; mean age 22.9, range 18-31) without psychiatric or neurological problems, with normal or corrected-to-normal vision and without hearing problems took part in the experiment. All participants except one (see Willems et al. 2014 for justification of inclusion) were right-handed by self-report, and all participants were naive with respect to the purpose of the experiment. Written informed consent was obtained in accordance with the Declaration of Helsinki, and the study was approved by the local ethics committee. Participants were paid either in money or in course credit at the end of the study.

Stimuli

Stimuli were taken from the Corpus of Spoken Dutch, ‘Corpus Gesproken Nederlands’ (Oostdijk, et al. 2000). Recordings were originally produced as part of the ‘Library for the Blind’, and comprised excerpts from three literary novels, all published in 1999 (Table 1). The excerpts were spoken at a normal rate, in a quiet room, by female speakers (one speaker per story). Stimuli durations were 3:49 min (622 words), 7:50 min (1291 words), and 7:48 min (1131 words). Reversed speech versions of the stories were created with Audacity 2.03 (http://audacity.sourceforge.net/). Descriptive statistics of the stories are displayed in Table 1.

		Word duration (ms)				**Lexical frequency**** (per million words)			
Stimulus*		**Mean**	**Median**	**Range**	**s.d.**	**Mean**	**Median**	**Range**	**s.d.**
Story 1	622	273	218	4-1174	181	5750	1539	0.02-39883	8306
Story 2	1291	252	193	31-949	160	6317	2106	0.02-39883	8876
Story 3	1131	274	212	40-1221	183	6612	1694	0.02-39883	9483

Table 1. Characteristics of the stimuli. Descriptive statistics for word duration and lexical frequency per story. S.d. = standard deviation.

*Story 1: from Peper, R., *Dooi,* L.J. Veen, 1999; Story 2: from Van der Meer, V., *Eilandgasten,* Contact, 1999; Story 3: from Jakobsen, A., *De Stalker*, De Boekerij, 1999

**Lexical frequency estimates were taken from the 44-million-word Subtlex NL database (Keuleers et al. 2010).

Estimation of surprisal and entropy

The conditional word probabilities required for obtaining surprisal and entropy values can be estimated by any probabilistic language model that is trained on a sufficiently large text corpus. We opted for a simple, efficient, and widely applied type of language model: the 2nd-order Markov model, more commonly known as *trigram model*. It is based on the simplifying assumption that the probability of word w_t depends on the previous two words *only*, that is, $P(w_t|w_1,\ldots,w_{t-1})$ is reduced to $P(w_t|w_{t-2},w_{t-1})$. Surprisal estimates by trigram models have been used successfully to account for experimental data from reading studies. For example, Frank et al. (2015) showed that trigram-based surprisal correlates positively with the N400 effect, and Smith and Levy (2013) found a linear relation with word reading time. Hence, previous research shows that the probabilities derived from trigram models accurately describe behavioural and neural indices of language comprehension.

One reason why trigram models are rather accurate is that the probabilities $P(w_t|w_{t-2},w_{t-1})$ can be reliably obtained from very large data sets. Here, we used a random selection of 10 million sentences (comprising 197 million word tokens; 2.1 million types) from the Dutch Corpus of Web (Schäfer and Bildhauer 2012). Based on this trigram model, for each word of the experimental texts, surprisal and entropy values were computed by the SRILM (Stolcke, 2002) and WOPR (Van den Bosch and Berck, 2009) software packages, respectively.

Occasionally, a stimulus word is not present in the training data, which means it receives a zero probability and, therefore, an infinite surprisal. These values were replaced by the largest finite value estimated for the three narratives, that is, unknown words are considered highly unlikely rather than impossible. This is equivalent to assuming the reasonable belief that any word has a non-zero probability of occurring, irrespective of the context.

Procedure

Participants listened to the three stories, as well as to the reversed speech versions of the stories, while they were lying in the MRI scanner. Each story and its reversed speech counterpart were presented following each other. Half of the participants started with a non-reversed stimulus, and half with a reversed speech stimulus. Participants were instructed to listen to the materials attentively. There was a short break after each fragment.

Stimuli were presented with Presentation software (version 16.2, http://www.neurobs.com). Auditory stimuli were presented through MR-compatible earphones. Presentation of the story fragments was preceded by a volume test: a fragment from another story with comparable voice and sound quality was presented while the scanner was collecting images. Volume was adjusted to the optimal level based on feedback from the participant.

Post-hoc memory test

After the scanning session participants were surprise-tested for their memory and comprehension of the stories. The post-hoc memory test was performed after all stories had been listened to. This was done with five multiple-choice questions per story fragment, with three answer options to each question. Questions were about general content, and memory scores were summed, leading to an overall score of each participant's memory of the story.

fMRI data acquisition and preprocessing

Images of Blood-Oxygenation Level Dependent (BOLD) changes were acquired on a 3T Siemens Magnetom Trio scanner (Erlangen, Germany) with a 32-channel head coil. Pillows and tape were used to minimize participants' head movement, and the earphones that were used for presenting the stories reduced scanner noise. Functional images were acquired using a fast T2*-weighted 3D EPI sequence (Poser et al. 2010), with high temporal resolution (TR: 880 ms, TE: 28 ms, flip angle: 14 degrees, voxel size: 3.5 x 3.5 x 3.5 mm, 36 slices). High resolution (1 x 1 x 1.25 mm) structural (anatomical) images were acquired using an MP-RAGE T1 GRAPPA sequence.

Preprocessing was performed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm) and Matlab 2010b (http://www.mathworks.nl/). After removing the first four volumes ('scans') to control for T1 equilibration effects, images were realigned to the first image in a run using rigid body registration ('motion correction'). The mean of the motion-corrected images was then brought into the same space as the individual participants' anatomical scan. The anatomical and functional scans were spatially normalized to the standard MNI template, and resampled to 2x2x2 mm voxel sizes. Finally, all data were spatially smoothed using an isotropic 8 mm full-width-at-half-maximum (FWHM) Gaussian kernel.

Data analysis

At the single-subject level, statistical analysis was performed using the general linear model, which means that the observed BOLD time course in each voxel is subjected to a regression analysis, testing for voxels in which the covariates of interest (surprisal and entropy) explain a significant proportion of variance of that voxel's time course (Friston et al. 1995). For each story, one regressor was created, modelling the duration of each single word. This regressor was convolved with the hemodynamic response function, to account for the delay in BOLD activation respective to stimulus presentation. Additionally three covariates (called 'parametric modulations' in SPM8) were added, one containing each word's log-transformed lexical frequency as determined from the Subtlex NL corpus (Keuleers et al.

2010), one containing each word's surprisal measure, and one containing the next-word entropy for each word. Log-transformed lexical frequency per word was added as a covariate of no interest, to statistically factor out effects of general word frequency, that is, expectations not based on linguistic context but on general word usage. Note that the entropy measure quantifies the uncertainty of the *upcoming* word, that is, the word at time *t*+1, whereas lexical frequency and word surprisal were taken for the word itself (the word at time *t*).

The same model was applied to the data from the reversed speech stimuli. That is, the word duration regressor and the covariates for a story were also fitted to the data of the reversed speech version of that story. The modelled time courses from all six runs (3 stories and 3 reversed speech stimuli) were combined in one regression model, with separate constant terms per run, but the same regressors for real and reversed speech. The estimates from the motion correction algorithm (3 rotations and 3 translations per run) were included in the model as regressors of no interest, to explain additional variance related to small head movements.

Whole-brain analysis involved group statistics in which participants were treated as a random factor (random effects analysis). The difference in the effect (i.e., regression slope) of the surprisal covariate and the entropy covariate between the real and reversed speech fragments for every voxel was used as input to the group level statistics. Statistical differences were assessed by computing the t-statistic over participants of this difference score (real versus reversed speech) for each voxel in the brain. The resulting multiple comparisons problem was solved by means of combining a $p<0.005$ voxel threshold with a cluster extent threshold determined by means of 2000 Monte Carlo simulations, after estimation of the smoothness of the data (Slotnick et al. 2003). This revealed that clusters of 54 contiguous voxels (resampled 2x2x2 mm voxels) or larger indicated statistically significant effects at the $p<0.05$ level, corrected for multiple comparisons. This cluster threshold was applied in all analyses.

Given our a priori hypothesis about inferior frontal cortex involvement, and given that whole brain analyses are necessarily conservative due to the correction for multiple comparisons, we additionally supplemented the whole brain analysis with region of interest (ROI) analyses. These were done using Marsbar (Brett et al. 2002), and comprised taking the mean contrast value per contrast of interest and per participant of all voxels in a ROI. Based on previous literature (see Introduction), we selected Brodmann Areas (BA) 44 and 45, both on the left and on the right side, resulting in four ROIs. The regions were defined using a cytoarchitectonic probability map (Amunts et al. 1999; Eickhoff et al. 2006).

Results

Behavioural

Participants answered on average 10.0 questions correctly (s.d. 2.21), out of 15 multiple choice questions asked, indicating memory performance above chance.

Whole-brain analysis

We first looked for regions that were sensitive to entropy (stronger negative relationship with entropy during real speech as compared to when listening to reversed speech). Activated regions include the right inferior frontal gyrus, the left ventral premotor cortex, extending into left middle frontal gyrus, the left supplementary motor area (SMA), and left inferior parietal lobule (Table 2; Figure 1).

Additionally we looked for voxels across the whole brain whose activation was more strongly modulated by surprisal during the real speech as compared to the reversed speech fragments. This means that activated regions were sensitive to surprisal (more surprisal leading to higher activation levels), and were more so in the real speech conditions as compared to the reversed speech conditions. Statistically significant activations were observed in the left inferior temporal sulcus / posterior fusiform gyrus, and left posterior superior temporal gyrus. Additionally there was an extensive activation spanning the right anterior temporal pole, right inferior frontal gyrus, right amygdala, and right brain stem. Finally there was a cluster of activation in the right posterior superior temporal gyrus (Table 3; Figure 1).

Region of interest analyses

As described above, we tested for effects of surprisal and entropy in four regions of interest: left BA44, right BA44, left BA45, and right BA45. No effects were observed for entropy (Table 4). Two-sided t-tests showed that in the right, but not in the left inferior frontal cortex, activation levels (beta weights) to surprisal in the non-reversed (story) condition were significantly different from the reversed-speech (control) condition (Table 4). In Figure 2 we plot the regression coefficients (beta weights) for surprisal and entropy, for each of the four ROIs.

Region	MNI coordinates (X Y Z)	Cluster extent (voxels)	Max t-value
Right inferior frontal gyrus	54 18 40	180	−3.99
	48 18 48		
Left middle frontal gyrus / ventral precentral sulcus	-46 12 50	353	−3.53
	-42 8 56		
	-36 30 46		
Left Supplementary Motor area	-4 42 44		−3.14
Left inferior parietal lobule	-44 -56 50	235	−3.50
	-52 -56 48		
	-46 -48 48		

Table 2. Entropy: results of whole-brain analysis. Areas activated to the entropy regressor. A negative relationship with entropy was tested, which means that activated regions had a stronger negative relationship with entropy during real than during reversed speech. Displayed are a description of the area, the MNI coordinates of the peak voxel, the cluster extent of the cluster, and the t-value of the peak voxel in the cluster. For larger clusters, several peak voxels' coordinates are provided to give a better representation of where activations were observed. All activations survived a p<0.05 FWE corrected statistical threshold.

Region	MNI coordinates (X Y Z)	Cluster extent (voxels)	Max t-value
L inferior temporal sulcus / posterior fusiform gyrus	-46 -46 -16	464	6.46
L posterior superior temporal gyrus	-64 -36 12	1345	5.84
	-58 -50 10		
	-54 -58 6		
L anterior temporal pole	-52 0 -4	117	3.47
	-56 8 -6		
R anterior temporal pole	56 8 -6	813	3.56
R hippocampus	30 -2 -11		3.50
R brain stem	12 -12 -8		3.71
R amygdala	32 -6 -12		3.11
R inferior frontal gyrus	50 12 -4		3.37
R superior temporal gyrus	64 -26 10	476	4.86

Table 3. Surprisal: results of whole-brain analysis. Areas activated to the surprisal regressor. Activated regions had a higher positive relationship with surprisal during real than during reversed speech. Displayed are a description of the area, the MNI coordinates of the peak voxel, the cluster extent of the cluster, and the t-value of the peak voxel in the cluster. For larger clusters, several peak voxels' coordinates are provided to give a better representation of where activations were observed. All activations survived a p<0.05 FWE corrected statistical threshold.

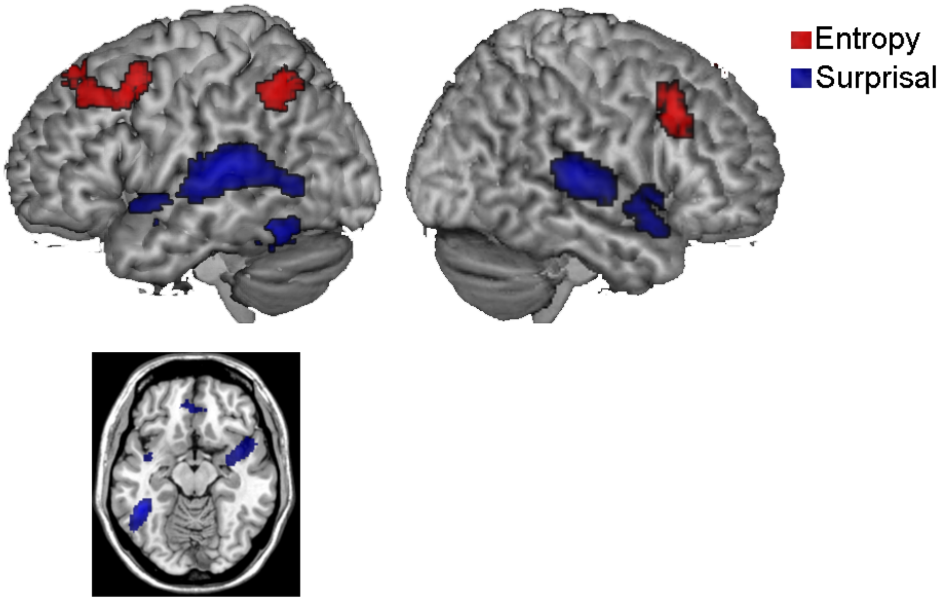

Figure 1. Results of the whole-brain analysis. Brain areas that were significantly activated in response to the regressors modelling entropy (red), and surprisal (blue) (see Tables 2 and 3). The inset shows the activation in hippocampus and amygdala (right hemisphere), and the activations in the inferior temporal and fusiform gyrus in response to word surprisal. Results are corrected for multiple comparisons at the p<0.05 FWE level.

	Entropy		Surprisal	
	t(23)	p-value	t(23)	p-value
L BA44	|t|<1	n.s.	1.18	0.25
R BA44	|t|<1	n.s.	**2.07**	**0.05**
L BA45	-1.31	0.20	1.47	0.16
R BA45	|t|<1	n.s.	**2.46**	**0.02**

Table 4. Results of the Region of interest analyses. The t-values indicate the difference in fit to surprisal or entropy between non-reversed ('real') and reversed speech. BA = Brodmann Area.

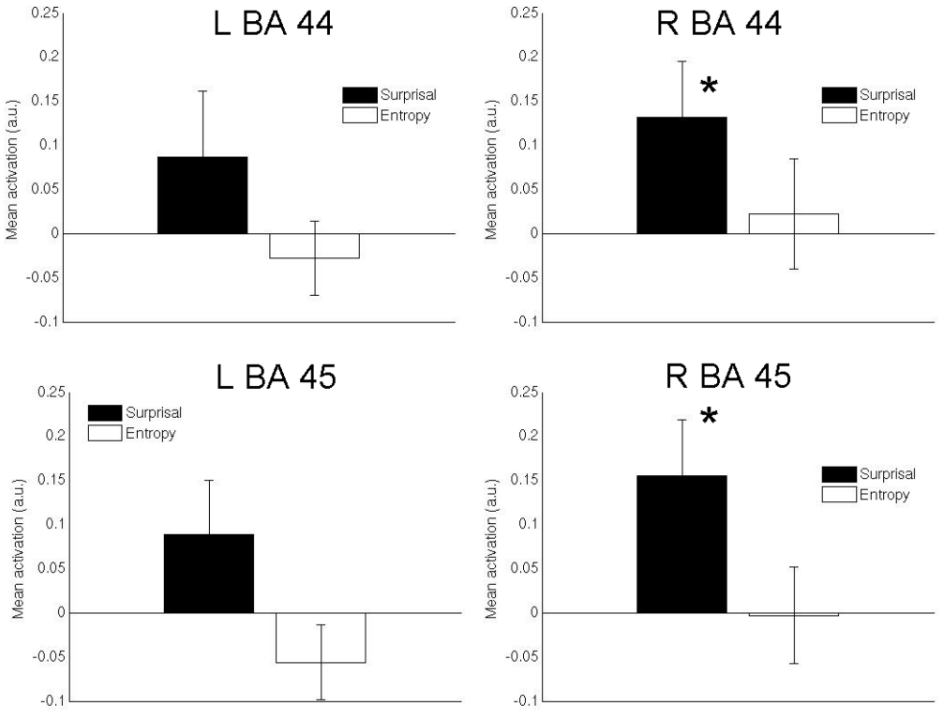

Figure 2. Results of the Regions of Interest analyses. Mean contrast values (Beta weights) in four regions of interest to word surprisal (black bars) and entropy (white bars). Error bars represent standard error of the mean (s.e.m.). Results show that right BA 44 and right BA 45 are sensitive to surprisal. This is somewhat at odds with the whole-brain analysis, in which the right inferior frontal gyrus (overlapping with right BA45) was sensitive to entropy. However, closer inspection reveals that the activation in the whole-brain analysis was rather specific to a subpart of the RIFG, and that the activation only comprised 5.1% of right BA45 (Eickhoff et al. 2005), which gets averaged out in the ROI analysis.

Discussion

We investigated the neural infrastructure related to prediction during natural language comprehension. A probabilistic language model was trained to predict upcoming words based upon co-occurrences in a 10-million-word corpus of written Dutch. This model was subsequently presented with three short narratives, and it provided two well-established information-theoretic measures for each word in the narratives. First, we looked at the entropy of the probability distribution of upcoming words. This measure indicates how uncertain the model is about the upcoming word. If the distribution of probabilities is broad, it means that many word candidates can be (weakly) predicted to be the next upcoming word, whereas if the distribution is narrow, only a limited subset of words can be (strongly) predicted. In other words, when entropy is low, the model is relatively certain about which word will be the next word, whereas when entropy is high, the model is less certain about the upcoming word. The second measure was surprisal, which is a quantification of the degree to which an incoming word was surprising, given its preceding context. The model compares the incoming word with its prediction before the word was perceived. Surprisal value is low when the actual perceived word was assigned a high occurrence probability, and is high when the perceived word was estimated to have low probability.

Participants' brain activation was measured with fMRI while they listened to spoken versions of the three narratives, as well as to the reversed speech versions of the stimuli (control stimuli). We found that entropy was negatively correlated with brain activation levels in left ventral premotor cortex, left middle frontal gyrus, right inferior frontal gyrus, left supplementary motor area, and left inferior parietal lobule. This means that these areas had higher activation when uncertainty about the upcoming word was low (i.e., predictability was high), and had lower activation levels when uncertainty was high (predictability was low). Surprisal value correlated significantly and positively with changes in brain activation levels in left inferior temporal sulcus / fusiform gyrus, bilateral anterior temporal poles, right amygdala / hippocampus, right inferior frontal sulcus, and posterior superior temporal gyri bilaterally. This means that these brain areas had increased activation when surprisal value was high, that is, when the perceived word deviated from what was predicted.

The interpretation of our results (see below) is consistent with hierarchical Bayesian inference as implemented by predictive coding in the brain (e.g. Friston 2005, Clark 2013). In this framework, predictions cancel prediction errors at lower levels of cortical hierarchies, where the precision or modulation of prediction errors depends upon the words' predictability. This is consistent with the activations related to entropy, such as middle frontal and ventral premotor cortex. A more surprising (i.e., less predicted) word generates prediction errors that are greater in amplitude and take longer to suppress. This we see reflected in the activations related to surprisal, for example, of the fusiform gyrus.

As explained in the Supplementary Materials, we also investigated trigram models that were trained on smaller text corpora, making them less accurate models of the language. In general, models that were trained on smaller corpora generated surprisal and entropy values that were less predictive of brain activation (Supplementary Figures S1 and S2). This relation between the quality of the language model and its fit to experimental data has been found before for reading time and N400 amplitude, in studies employing single sentence reading (Monsalve et al. 2012; Frank 2013; Frank et al. 2015). This is a relevant finding because it adds plausibility to the conclusion that the effects of surprisal and entropy on brain activation are due to the language model's predictions and not to some other unrelated factor.

Activations related to entropy

Left ventral premotor cortex showed a positive relationship with predictability: When predictability of the upcoming word was high (i.e., entropy was low), activation in this area was high as well. The ventral premotor cortex has been found to be sensitive to entropy before. For instance, Nastase and colleagues presented participants with series of auditory and visual stimuli, which differed in entropy from completely random to highly ordered sequences. The auditory stimuli were pure tones, and the visual stimuli consisted of simple coloured shapes (e.g. a blue triangle) (Nastase et al. 2014). The left ventral premotor cortex (among other regions) was sensitive to the level of entropy in the series in both modalities. Similarly, Schubotz and Von Cramon (2004) found the left ventral premotor cortex to be sensitive to predictability across different modalities. They presented participants with sequences of actions (e.g. someone putting a paper into a post box), and also with sequences of abstract shapes, and found left ventral premotor cortex to be sensitive to predictability of both types of stimuli. Nastase and colleagues (2014) also observed sensitivity of the supplementary motor area (SMA) to entropy, a result which we replicate here.

Another brain structure which has been implicated into the encoding of entropy is the hippocampus (e.g. Strange et al. 2005). This was not replicated in the present study, instead the hippocampus was found to be activated in response to word surprisal (see below).

What we labelled inferior parietal lobule contains the angular gyrus. Binder and colleagues concluded in their extensive meta-analysis that this area 'occupies a position at the top of a processing hierarchy underlying concept retrieval and conceptual integration' (Binder et al. 2009, p. 2776). Indeed, the angular gyrus is considered a hub in the language network (Turken and Dronkers 2011), and our results are compatible with the claim that it is involved in the active prediction of upcoming words.

The left middle frontal gyrus and right inferior frontal cortices were also sensitive to entropy. It is tempting to interpret this activation as a top-down influence from these areas onto parts of the language network lower in the cortical hierarchy. When predictability is high, left middle frontal and right inferior frontal cortex can presumably pre-activate representations in other areas. These areas lend themselves well for such a modulatory function, as has been observed in previous language-related research (Fiebach et al. 2006; Snijders et al. 2010).

Activations related to surprisal

A cluster of activation in the left inferior temporal sulcus / fusiform gyrus was found to be sensitive to surprisal. This area, which is a well-known activation site in studies of language, is sensitive to word form, albeit that parts of this area are also sensitive to lexical-semantic features of words (Vinckier et al. 2007; Levy et al. 2009). Comparing presentation of words versus non-words for instance activates this area, which has been dubbed 'visual word form area' (VWFA; Cohen et al. 2000). Activation of the VWFA in the present experiment with auditory presentation may seem surprising. However, it should be noted that parts of the inferior temporal sulcus, neighbouring the VWFA, and overlapping with the present activation, have been found to be sensitive to both written and auditory word forms (Cohen et al. 2004). Cohen and colleagues dubbed this part of inferior temporal cortex the 'lateral inferotemporal multimodal area' (LIMA), an acronym that did not become nearly as popular as VWFA. The sensitivity of this region to surprisal suggests that during natural language comprehension, there might be priming of predicted word forms, which allows the system to quickly process incoming information. Conversely, when the incoming word violates the prediction (high word surprisal), this results in prediction error. This result fits well with an EEG study showing that left posterior electrodes (presumably reflecting activation of the word form area) distinguish predictable versus less predictable words in a multi-sentence context within 90 milliseconds after reading of the word (Dambacher et al. 2009). Such an early effect argues for a role of prediction, which is in line with our current interpretation.

A similar interpretation is to be given to the bilateral activations of the superior temporal areas that we observed. The superior temporal gyrus comprises the primary and secondary auditory cortices, and here we show that these areas are modulated by how well an incoming word fits the predicted input. This is reminiscent of recent findings concerning the influence of prediction during visual perception. In those cases too, early visual areas are sensitive to the content of the predicted visual stimulus (e.g. Kok et al. 2012). Here we extend the early role of prediction to natural language comprehension, by showing that early auditory areas, as well as areas coding for word forms, are sensitive to the surprisal of an incoming word (see Dikker et al. 2010, 2014; Molinaro et al. 2013).

A related low-level activation was observed in the right amygdala and part of the hippocampus. It is tempting to interpret the sensitivity of the right amygdala to word surprisal in the context of the intracranial recordings by Nobre and McCarthy and colleagues. They observed differences between normal and anomalous sentence endings (the N400 paradigm) at several electrode sites close to the amygdala. Intracranial electrodes in the amygdala were sensitive to the manipulation, but based on simultaneous local field recordings they conclude that the neural generator of the observed difference between anomalous and correct sentence endings is not the amygdala itself but the nearby portion of the anterior fusiform gyrus (McCarthy et al. 1995; see also Nobre et al. 1994). In other language studies the amygdala has been mostly linked to processing of emotional content of words (e.g. Herbert et al. 2009; Willems et al. 2011; Chow et al. 2014), and here we show that this part of the brain may also be sensitive to a word's expectedness. Converging evidence for the role of the amygdala in prediction during language comprehension comes from a recent study in which it was found that activation in the right amygdala was higher just before participants heard a highly expected word (J. Skipper, personal communication).

The bilateral temporal poles' sensitivity to surprisal is best understood with regard to the integrating function which has been ascribed to the anterior temporal poles (e.g. Brennan et al. 2012). Activation of the anterior temporal poles is often observed when participants comprehend narratives (e.g. Mazoyer et al. 1993; Vandenberghe et al. 2002; Crinion et al. 2003, 2006; Xu et al. 2005; Awad et al. 2007; Ferstl et al. 2008), and their function is thought to be related to forming a unified whole within and across sentences. Moreover, Lau and colleagues showed that the anterior temporal poles are sensitive to the extent to which perceived word pairs can be predicted from the context (Lau et al. 2014). The fact that the anterior temporal poles showed sensitivity to word surprisal values adds to the psychological plausibility of word surprisal.

Finally, we observed that right inferior frontal cortex was sensitive to word surprisal. The right IFG has been found sensitive to narrative and discourse comprehension (e.g. St George et al. 1999; Robertson et al. 2000; Xu et al. 2005), although the consistency of right IFG activation during discourse comprehension is debated (Ferstl et al. 2008; Hagoort et al. 2009). The double role of the right inferior frontal cortex in this study (sensitive to both entropy and surprisal) could reflect distinguishable subfunctions of parts of this region (see Figure 2), and this deserves further attention in future studies.

Interestingly, we found no modulatory effects of word surprisal on several areas which have been implicated in discourse or narrative comprehension. For instance, the anterior medial prefrontal cortex as well as the posterior cingulate cortex / precuneus are found to be more activate during narrative comprehension than during comprehension of single, unstructured sentences (Lerner et al. 2011). One reason for the insensitivity to surprisal could be that the surprisal value as operationalised here is sensitive to local context, and not to global context spanning for instance several paragraphs or even a single sentence (but see Tobia et al. 2012). Also left inferior frontal cortex was not sensitive to word surprisal, despite its well-established role in language comprehension. It is hard to interpret this absence, given that it could be due to a lack of statistical power. Figure 2B indeed shows that left and right inferior frontal cortex show a similar pattern of responses, but that only in right inferior frontal cortex the response is robust enough to reach statistical significance. We therefore refrain from drawing strong conclusions about hemispheric differences between the left and right inferior frontal areas based on this result alone.

Conclusion

We investigated the neural implementation of entropy and surprisal of words during natural language comprehension. Left ventral premotor cortex was sensitive to entropy of the probability distribution of the upcoming word, as has been found before for non-language stimuli. Similarly, an important node in the semantic brain network, the left inferior parietal lobule ('angular gyrus') also showed sensitivity to entropy. Together with activations in left middle frontal and right inferior frontal cortex, that were also sensitive to entropy, these are areas that most likely exhibit a top-down influence related to predictability of the upcoming word, onto other areas in the language network. Additionally, we found that areas relatively early in the neural language network are sensitive to surprisal, that is, how predictable the current word was given the previous context. The modulations of left inferior temporal sulcus, bilateral posterior superior temporal gyri, and right amygdala point to modulating effects of prediction onto lower levels of processing during language comprehension. This suggests that prediction can occur early in the processing stream, already at the level of word form (see Dikker et al. 2010, 2014; Molinaro et al. 2013). Anterior temporal poles and right inferior frontal cortex are similarly sensitive to word surprisal, and, based on previous literature, activation of these areas is most likely best explained as a combination of predictive and integrative functions (Brennan and Pylkkänen 2012; Hagoort 2013).

Besides the exact interpretation of functional localization in our results, the current study shows that it is fruitful to use well-defined information-theoretic measures from computational linguistics to inform us about the neural basis of natural language comprehension. A methodological advance of the present study is that participants listened to relatively long stretches of natural language, and that characterization of the stimuli was done in a computationally explicit manner. A growing body of research into the neural basis of language uses more naturalistic stimuli (Nijhof and Willems, 2015; Speer et al. 2009; Lerner et al. 2011; Wallentin et al. 2011; Brennan et al. 2012; Kurby and Zacks 2013; Altmann et al. 2014; Hsu et al. 2014; Wehbe et al., 2014; Jacobs 2015). This is more than a methodological nicety: In several fields of cognitive neuroscience, researchers suggest that knowledge derived using highly simplified and context-free stimuli does not always generalize to more natural situations (e.g. Olshausen and Field 2005; Ferreira and Patson 2007; Peelen and Kastner 2014; Willems

2015). Our study provides insights into how prediction affects the language system under relatively natural circumstances of language comprehension.

Acknowledgements

We are thankful to Markus Barth and Paul Gaalman for assistance in MRI sequence selection. Two anonymous reviewers provided insightful suggestions to an earlier version of this manuscript, for which we are most thankful. The research presented here was partly funded by the European Union Seventh Framework Programme under grant number 334028 awarded to SLF.

References

Ahlheim C, Stadler W, Schubotz RI. 2014. Dissociating dynamic probability and predictability in observed actions—an fMRI study. Front Hum Neurosci. 8.

Altmann U, Bohrn IC, Lubrich O, Menninghaus W, Jacobs AM. 2014. Fact vs fiction--how paratextual information shapes our reading processes. Soc Cogn Affect Neurosci. 9:22–29.

Amunts K, Schleicher A, Burgel U, Mohlberg H, Uylings HB, Zilles K. 1999. Broca’s region revisited: cytoarchitecture and intersubject variability. J Comp Neurol. 412:319–341.

Awad M, Warren JE, Scott SK, Turkheimer FE, Wise RJS. 2007. A common system for the comprehension and production of narrative speech. J Neurosci Off J Soc Neurosci. 27:11455–11464.

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

Binder JR, Desai RH, Graves WW, Conant LL. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. Cereb Cortex N Y N 1991. 19:2767–2796.

Brennan J, Nir Y, Hasson U, Malach R, Heeger DJ, Pylkkänen L. 2012. Syntactic structure building in the anterior temporal lobe during natural story listening. Brain Lang. 120:163–173.

Brennan J, Pylkkänen L. 2012. The time-course and spatial distribution of brain activity associated with sentence processing. NeuroImage. 60:1139–1148.

Brett M, Anton J-L, Valabregue R, Poline J-B. 2002. Region of interest analysis using an SPM toolbox. Neuroimage. 16.

Chen SF, Goodman J. 1999. An empirical study of smoothing techniques for language modeling. Comput Speech Lang. 13:359–394.

Chow HM, Mar RA, Xu Y, Liu S, Wagage S, Braun AR. 2014. Embodied comprehension of stories: interactions between language regions and modality-specific neural systems. J Cogn Neurosci. 26:279–295.

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

Cohen L, Dehaene S, Naccache L, Lehericy S, Dehaene-Lambertz G, Henaff MA, Michel F. 2000. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain. 123 (Pt 2):291–307.

Cohen L, Jobert A, Le Bihan D, Dehaene S. 2004. Distinct unimodal and multimodal regions for word processing in the left temporal cortex. Neuroimage. 23:1256–1270.

Crinion JT, Lambon-Ralph MA, Warburton EA, Howard D, Wise RJS. 2003. Temporal lobe regions engaged during normal speech comprehension. Brain J Neurol. 126:1193–1201.

Crinion JT, Warburton EA, Lambon-Ralph MA, Howard D, Wise RJS. 2006. Listening to narrative speech after aphasic stroke: the role of the left anterior temporal lobe. Cereb Cortex N Y N 1991. 16:1116–1125.

Dambacher M, Rolfs M, Göllner K, Kliegl R, Jacobs AM. 2009. Event-Related Potentials Reveal Rapid Verification of Predicted Visual Input. PLoS ONE. 4:e5047.

DeLong KA, Urbach TP, Kutas M. 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat Neurosci. 8:1117–1121.

Den Ouden HEM, Kok P, de Lange FP. 2012. How prediction errors shape perception, attention, and motivation. Front Psychol. 3:548.

Dikker S, Pylkkänen L. 2013. Predicting language: MEG evidence for lexical preactivation. Brain Lang. 127:55–64.

Dikker S, Rabagliati H, Farmer TA, Pylkkänen L. 2010. Early occipital sensitivity to syntactic category is based on form typicality. Psychol Sci. 21:629–634.

Dikker S, Silbert LJ, Hasson U, Zevin JD. 2014. On the same wavelength: predictable language enhances speaker-listener brain-to-brain synchrony in posterior superior temporal gyrus. J Neurosci Off J Soc Neurosci. 34:6267–6272.

Eickhoff SB, Heim S, Zilles K, Amunts K. 2006. Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. Neuroimage. 32:570–582.

Federmeier KD. 2007. Thinking ahead: the role and roots of prediction in language comprehension. Psychophysiology. 44:491–505.

Ferreira F, Patson ND. 2007. The “Good Enough” Approach to Language Comprehension. Lang Linguist Compass. 1:71–83.

Ferstl EC, Neumann J, Bogler C, von Cramon DY. 2008. The extended language network: A meta-analysis of neuroimaging studies on text comprehension. Hum Brain Mapp. 29:581–593.

Fiebach CJ, Rissman J, D’Esposito M. 2006. Modulation of inferotemporal cortex activation during verbal working memory maintenance. Neuron. 51:251–261.

Frank SL. 2013. Uncertainty reduction as a measure of cognitive load in sentence comprehension. Top Cogn Sci. 5:475–494.

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

Frank SL, Thompson RL. 2012. Early effects of word surprisal on pupil size during reading. In: Proceedings of the 34th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society. p. 1554–1559.

Friston K. 2005. A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci. 360:815–836.

Friston KJ, Holmes A, Worsley KJ, Poline J-B, Frith CD, Frackowiak RS. 1995. Statistical Parametric Maps in Functional Imaging: A General Linear Approach. Hum Brain Mapp. 2:189–210.

Hagoort P. 2005. On Broca, brain, and binding: a new framework. Trends Cogn Sci. 9:416–423.

Hagoort P. 2013. MUC (Memory, Unification, Control) and beyond. Front Psychol. 4:416.

Hagoort P, Baggio G, Willems RM. 2009. Semantic unification. In: Gazzaniga MS, editor. The cognitive neurosciences IV. MIT press.

Hale JT. 2001. A probabilistic Early parser as a psycholinguistic model. In: Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics. Pittsburgh, PA: Association for Computational Linguistics. p. 159–166.

Herbert C, Ethofer T, Anders S, Junghofer M, Wildgruber D, Grodd W, Kissler J. 2009. Amygdala activation during reading of emotional adjectives--an advantage for pleasant content. Soc Cogn Affect Neurosci. 4:35–49.

Hsu C-T, Jacobs AM, Conrad M. 2014. Can Harry Potter still put a spell on us in a second language? An fMRI study on reading emotion-laden literature in late bilinguals. Cortex J Devoted Study Nerv Syst Behav. 63C:282–295.

Jacobs AM. 2015. Towards a Neurocognitive Poetics Model of Literary Reading. In: Willems RM, editor. Cognitive Neuroscience of Natural Language Use. Cambridge, UK: Cambridge University Press.

Keuleers E, Brysbaert M, New B. 2010. SUBTLEX-NL: a new measure for Dutch word frequency based on film subtitles. Behav Res Methods. 42:643–650.

Koehn P. 2010. Statistical machine translation. Cambridge, UK: Cambridge University Press.

Kok P, Jehee JFM, de Lange FP. 2012. Less is more: expectation sharpens representations in the primary visual cortex. Neuron. 75:265–270.

Kurby CA, Zacks JM. 2013. The activation of modality-specific representations during discourse processing. Brain Lang. 126:338–349.

Kutas M, Federmeier KD. 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu Rev Psychol. 62:621–647.

Laszlo S, Federmeier KD. 2009. A Beautiful Day in the Neighborhood: An Event-Related Potential Study of Lexical Relationships and Prediction in Context. J Mem Lang. 61:326–338.

Lau EF, Weber K, Gramfort A, Hämäläinen MS, Kuperberg GR. 2014. Spatiotemporal Signatures of Lexical-Semantic Prediction. Cereb Cortex N Y N 1991.

Lerner Y, Honey CJ, Silbert LJ, Hasson U. 2011. Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J Neurosci Off J Soc Neurosci. 31:2906–2915.

Levy J, Pernet C, Treserras S, Boulanouar K, Aubry F, Démonet J-F, Celsis P. 2009. Testing for the dual-route cascade reading model in the brain: an fMRI effective connectivity account of an efficient reading style.PloS One. 4:e6675.

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

Mahowald K, Fedorenko E, Piantadosi ST, Gibson E. 2013. Info/information theory: speakers choose shorter words in predictive contexts. Cognition. 126:313–318.

Mason RA, Just MA. 2004. How the Brain Processes Causal Inferences in Text A Theoretical Account of Generation and Integration Component Processes Utilizing Both Cerebral Hemispheres. Psychol Sci. 15:1–7.

Mazoyer BM, 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 CognNeurosci. 5:467–479.

McCarthy G, Nobre AC, Bentin S, Spencer DD. 1995. Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. J Neurosci Off J Soc Neurosci. 15:1080–1089.

Menenti L, Petersson KM, Scheeringa R, Hagoort P. 2009. When elephants fly: differential sensitivity of right and left inferior frontal gyri to discourse and world knowledge. J Cogn Neurosci. 21:2358–2368.

Molinaro N, Barraza P, Carreiras M. 2013. Long-range neural synchronization supports fast and efficient reading: EEG correlates of processing expected words in sentences. NeuroImage. 72:120–132.

Monsalve I. F., Frank SL, Vigliocco G. 2012. Lexical surprisal as a general predictor of reading time. In: Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics. Presented at the Association for Computational Linguistics. Avignon, France. p. 398–408.

Nastase S, Iacovella V, Hasson U. 2014. Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. Hum Brain Mapp. 35:1111–1128.

Nijhof AD, Willems RM. 2015. Simulating fiction: Individual differences in literature comprehension revealed with fMRI. PloS One. 10:e0116492

Nobre AC, Allison T, McCarthy G. 1994. Word recognition in the human inferior temporal lobe.Nature. 372:260–263.

Ogawa S, Lee TM, Kay AR, Tank DW. 1990. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. ProcNatlAcadSci U S A. 87:9868–9872.

Olshausen BA, Field DJ. 2005. How close are we to understanding v1? Neural Comput. 17:1665–1699.

Oostdijk, N.H.J., Gravilidou, M., Carayannis, S., Markantonatou, S., Piperidis, S., Stainhaouer, G. 2000. The spoken Dutch Corpus. Outline and first evaluation.Proc Second Int Conf Lang Resour Eval. 2:887–894.

Peelen MV, Kastner S. 2014. Attention in the real world: toward understanding its neural basis. Trends Cogn Sci.

Piantadosi ST, Tily H, Gibson E. 2011. Word lengths are optimized for efficient communication. Proc Natl Acad Sci U S A. 108:3526–3529.

Poser BA, Koopmans PJ, Witzel T, Wald LL, Barth M. 2010.Three dimensional echo-planar imaging at 7 Tesla.NeuroImage. 51:261–266.

Roark B,Bachrach A, Cardenas C, Pallier C. 2009. Deriving lexical and syntactic expectation-basedmeasures for psycholinguistic modeling viaincremental top-down parsing. Proc 2009 Conf Emp Meth Nat Lang Proc. p. 324–333

Robertson DA, Gernsbacher MA, Guidotti SJ, Robertson RRW, Irwin W, Mock BJ, Campana ME. 2000. Functional Neuroanatomy of the Cognitive Process of Mapping During Discourse Comprehension. Psychol Sci. 11:255–260.

Schäfer R, Bildhauer F. 2012.Building large corpora from the web using a new efficient tool chain. In: Proceedings of the 8th international conference on Language Resources and Evaluation. European Language Resources Association.

Schubotz RI, von Cramon DY. 2004. Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. J Neurosci. 24:5467–5474.

Slotnick SD, Moo LR, Segal JB, Hart J Jr. 2003. Distinct prefrontal cortex activity associated with item memory and source memory for visual shapes. Brain Res Cogn Brain Res. 17:75–82.

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

Snijders TM, Petersson KM, Hagoort P. 2010. Effective connectivity of cortical and subcortical regions during unification of sentence structure.NeuroImage. 52:1633–1644.

Speer NK, Reynolds JR, Swallow KM, Zacks JM. 2009. Reading stories activates neural representations of visual and motor experiences. Psychol Sci. 20:989–999.

St George M, Kutas M, Martinez A, Sereno MI. 1999. Semantic integration in reading: engagement of the right hemisphere during discourse processing. Brain. 122 (Pt 7):1317–1325.

Stolcke A. 2002. SRILM – an extensible language modeling toolkit. In: Proc Internat Conf Spoken Lang Proc. Denver, Colorado. p. 901–904.

Strange BA, Duggins A, Penny W, Dolan RJ, Friston KJ. 2005. Information theory, novelty and hippocampal responses: unpredicted or unpredictable? Neural Netw. 18:225–230.

Tobia MJ, Iacovella V, Davis B, Hasson U. 2012.Neural systems mediating recognition of changes in statistical regularities.NeuroImage. 63:1730–1742.

Turken AU, Dronkers NF. 2011. The Neural Architecture of the Language Comprehension Network: Converging Evidence from Lesion and Connectivity Analyses. Front Syst Neurosci. 5.

Van Berkum JJA, Brown CM, Zwitserlood 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 AC, Price CJ. 2002. The response of left temporal cortex to sentences. J Cogn Neurosci. 14:550–560.

Van den Bosch, A., Berck, P. 2009. Memory-based machine translation and language modeling. Prague Bull Math Ling 91:17-26.

Van Petten C, Luka BJ. 2012. Prediction during language comprehension: benefits, costs, and ERP components. Int J PsychophysiolOff J Int Organ Psychophysiol. 83:176–190.

Vinckier F, Dehaene S, Jobert A, Dubus JP, Sigman M, Cohen L. 2007. Hierarchical coding of letter strings in the ventral stream: dissecting the inner organization of the visual word-form system. Neuron. 55:143–156.

Wallentin M, Nielsen AH, Vuust P, Dohn A, Roepstorff A, Lund TE. 2011. BOLD response to motion verbs in left posterior middle temporal gyrus during story comprehension. Brain Lang. 119:221–225.

Wehbe L, Murphy B, Talukdar P, Fyshe A, Ramdas A, Mitchell T. 2014. Simultaneously uncovering the patterns of brain regions involved in different story reading subprocesses. PloS One. 9:e112575.

Wicha NYY, Moreno EM, Kutas M. 2004.Anticipating words and their gender: an event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. J Cogn Neurosci. 16:1272–1288.

Willems RM (Ed.). 2015. Cognitive Neuroscience of Natural Language Use. Cambridge, UK: Cambridge University Press.

Willems RM, Clevis K, Hagoort P. 2011. Add a picture for suspense: neural correlates of the interaction between language and visual information in the perception of fear. Soc Cogn Affect Neurosci. 6:404–416.

Willems RM, der Haegen LV, Fisher SE, Francks C. 2014. On the other hand: including left-handers in cognitive neuroscience and neurogenetics. Nat Rev Neurosci. 15:193–201.

Xu J, Kemeny S, Park G, Frattali C, Braun A. 2005. Language in context: emergent features of word, sentence, and narrative comprehension. Neuroimage. 25:1002–1015.