

# Multi-voxel pattern analysis of noun and verb differences in ventral temporal cortex



Christine Boylan\*, John C. Trueswell, Sharon L. Thompson-Schill

Department of Psychology, University of Pennsylvania, USA

## ARTICLE INFO

### Article history:

Accepted 27 July 2014

Available online 24 August 2014

### Keywords:

Multi-voxel pattern analysis

Prediction

Visual word form area

Syntax

Phonology

## ABSTRACT

Recent evidence suggests a probabilistic relationship exists between the phonological/orthographic form of a word and its lexical-syntactic category (specifically nouns vs. verbs) such that syntactic prediction may elicit form-based estimates in sensory cortex. We tested this hypothesis by conducting multi-voxel pattern analysis (MVPA) of fMRI data from early visual cortex (EVC), left ventral temporal (VT) cortex, and a subregion of the latter – the left mid fusiform gyrus (mid FG), sometimes called the “visual word form area.” Crucially, we examined only those volumes sampled when subjects were predicting, *but not viewing*, nouns and verbs. This allowed us to investigate prediction effects in visual areas without any bottom-up orthographic input. We found that voxels in VT and mid FG, but not in EVC, were able to classify noun-predictive trials vs. verb-predictive trials in sentence contexts, suggesting that sentence-level predictions are sufficient to generate word form-based estimates in visual areas.

© 2014 Elsevier Inc. All rights reserved.

## 1. Introduction

Language, like any other temporally ordered behavior, makes use of top-down predictions in order to reduce uncertainty about upcoming events. The fact that language processing is so remarkably fast is likely due to our ability to predict the types of structures found in natural language, whether these be phonological, morpho-syntactic, lexical-semantic, or pragmatic. Given the immense generative power of language, it is unlikely that linguistic prediction operates only over the surface statistics of a language; rather, efficiency would dictate that predictions be based on the language’s “category statistics,” or the likelihood that one set of elements is followed by another (Hunt & Aslin, 2010). The existence of linguistic categories such as, say, nouns and verbs, is relatively easy to determine, but the predictive power of these categories is limited, if not entirely obfuscated, by the apparently arbitrary relationship between a word’s syntactic category and the phonological features of that category’s members. The venerable principle of the “arbitrariness of the sign” has provided not only a descriptive account of why the phonological similarity of words, such as *cat*, *sat*, and *fat*, determines neither their semantic meaning nor syntactic category (de Saussure, 1916; Tanenhaus & Hare, 2007) but also a functional account: if a word’s form is uncoupled

from its meaning, this allows a finite set of forms to combine to denote an infinite set of meanings (Chomsky, 1965). Thus it would seem that language’s infinite generativity is at odds with optimal conditions for word form prediction.

However, a study by Farmer, Christiansen, and Monaghan (2006) provided evidence that a probabilistic relationship may indeed exist between the phonological/orthographic form of a word and its lexical category, which could in principle be used by a reader/listener to predict word form features during sentence processing. The study was prompted by a renewed interest in research demonstrating that systematic, probabilistic, form-based regularities exist among the words of a given lexical category (Arciuli & Monaghan, 2009; Cassidy & Kelly, 1991; Kelly, 1992; Monaghan, Christiansen, & Chater, 2007; but cf. Staub, Grant, Clifton, & Rayner, 2009). In a corpus analysis of the phonological properties of nouns and verbs, Farmer et al. found these two lexical categories formed distinct clusters when plotted in a multidimensional form feature space. They calculated the form feature distance between each possible two-word comparison based on the number of overlapping and non-overlapping phonetic features. They then obtained a “form typicality score” for each word by subtracting its distance to all verbs from its distance to all nouns. While many words were “neutral” – not strongly typical of either nouns or verbs – the centers of noun-typicality and verb-typicality were separated in this feature space such that clusters of typical nouns and typical verbs could be distinguished. Furthermore, the noun- or verb-typicality of a word was found to predict lexical naming latencies and reading times. This typicality measure also

\* Corresponding author. Present address: Institute for Research in Cognitive Science, University of Pennsylvania, 3401 Walnut Street, Suite 400A, Philadelphia, PA 19104, USA. Fax: +1 215 573 9247.

E-mail address: [chboylan@sas.upenn.edu](mailto:chboylan@sas.upenn.edu) (C. Boylan).

influenced syntactic processing: whether a noun–verb homonym was more typical of a noun or a verb predicted whether participants expected a noun or verb continuation of a given ambiguous sentence. The effect of this typicality measure was significant even after accounting for effects of onset phoneme, frequency, length, neighborhood size, familiarity, and imageability.

The present work uses fMRI multi-voxel pattern classification to test whether readers predict word forms corresponding to noun and verb syntactic categories and to examine the neural instantiation of these putative predictions. There are several candidates for the neural read-out of such a predictive system. In this study, we will explore areas where this prediction may engage the brain's extended visual system. Although Farmer et al. (2006) quantified form typicality using a phonological feature metric and not a visual orthographic metric per se, they found evidence that this form typicality metric predicted reading times. English's use of a phonemic orthography (in which graphemes have a correspondence to phonemes) leads one to expect that a syntactic–phonological–orthographic correspondence could play a role in using lexical category expectations to predict visual word form features. If so, we would expect such prediction to recruit areas of the brain sensitive to features of words and letter strings. One such candidate region is the left mid fusiform gyrus, referred to by some as the “visual word form area” due to its putative specialization in identifying visual word forms (Dehaene & Cohen, 2011). Although the functional specificity of this area is not uncontroversial, and there may be other areas of the brain subserving written word recognition, the left mid fusiform gyrus is robustly sensitive to visual word stimuli, and thus could be involved in generating word form predictions. We also looked at a larger swathe of ventral temporal cortex surrounding mid FG, since the mid FG may be part of a more diffuse posterior-to-anterior tuning gradient extending along the left ventral temporal cortex and sensitive to (non-)orthographic line junctions, alphabetic letters, bigrams, morphemes, and whole words (Haushofer, Livingstone, & Kanwisher, 2008; Vinckier et al., 2007).

Rather more controversial, however, is evidence that syntactic predictions during reading may generate form-based estimates as early as occipital cortex (Dikker, Rabagliati, Farmer, & Pyllkanen, 2010). In an event-related magnetoencephalographic (MEG) study, Dikker et al. compared brain responses across two syntactic violation conditions. In both conditions, the syntax of the sentence selected for a verb, but in one case the next word was a form-typical noun and in the other it was a form-neutral noun, which had form features consistent with both nouns and verbs. It was found that the amplitude of the MEG component called the M100 (i.e., 100 ms post-stimulus onset) was significantly greater when a typical noun violated the sentence continuation than when a neutral noun did. Although the type of syntactic violation was equivalent in both cases, only the typicality scores predicted this M100 modulation. In other work, the M100 has been localized to early visual cortex (EVC) – specifically the cuneus, lingual gyrus, and BA 17 (Itier, Herdman, George, Cheyne, & Taylor, 2006). Thus, these data compelled us to look at EVC in addition to more anterior regions in VT.

In the present work, we were concerned not only with the question of *where* in the brain lexical-syntactic categories might map onto form features, but also the questions of *how* and *when*. Could the early visual form typicality effect in MEG have marked an *in situ* violation detection, or might a lexical class violation generate an error signal elsewhere in the brain that is then relayed to visual areas via re-entrant pathways? Is the expectation violation detected first in higher-level areas, *after* the word has been fully analyzed for lexical syntactic properties, or do visual areas have enough information about the predicted word form features to “raise the first alarm”? One hypothesis entails top-down

prediction, while the other requires no such prediction, but rather a fast bottom-up analysis of a word before the lexical class violation can be detected. While the M100 has been shown to be sensitive to orthographic frequency and transition probability of letter strings, there is no evidence that the M100 is sensitive to lexical factors of words in isolation (Solomyak & Marantz, 2009; Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999). For this reason, the MEG findings led to the hypothesis that top-down prediction must be involved (Dikker et al., 2010); however, this hypothesis has not been directly tested until now.

One way to distinguish top-down prediction effects of word form estimation from bottom-up perceptual effects of word recognition is simply to remove the word stimulus. We did just this in the following experiment: we presented subjects with syntactically predictive sentence fragment cues followed by a series of random dot patterns in which the subject was to search for either a noun or a verb. Subjects viewed sentence fragments that highly constrained the category of word that could continue the sentence (e.g., a noun was expected but not a verb, or vice versa) but did not constrain expectation for a specific word within that category (see Appendix for list of stimuli). Instead of seeing the sentence-final word immediately, subjects searched for an appropriate sentence completion in a series of noisy images and indicated when an appropriate word was discernible (see Fig. 1).

There is precedent in the MVPA literature for successful decoding of imagined shapes (Stokes, Thompson, Cusack, & Duncan, 2009; Stokes, Thompson, Nobre, & Duncan, 2009), objects (Lee, Kravitz, & Baker, 2012), and object categories (e.g., people vs. cars; Peelen & Kastner, 2011) from distributed BOLD activity. Extending this method to highly abstract grammatical word categories, we were able to successfully classify nouns vs. verbs in VT and mid FG when a syntactic context was provided. In contrast, EVC did not support classification in this study. These results suggest that syntactic, or at least sentence-level, prediction prompts form-based estimates in early visual word form areas, and that a probabilistic relationship between word form and word category is indeed exploited by the neural circuitry.

## 2. Material and methods

### 2.1. Participants

Twelve subjects participated in this study. Two subjects' data were excluded due to excessive motion artifact, leaving ten subjects analyzed here. Subjects ranged in age from 18 to 38 years, and all were right-handed native speakers of English with normal or corrected-to-normal vision and no reported history of neurologic problems. Subjects gave written informed consent and were provided monetary compensation for their time. The human subjects review board at the University of Pennsylvania approved all experimental procedures.

### 2.2. Task and stimuli

#### 2.2.1. Sentence norming

The sentences used in this study were constructed such that the final word in the sentence could be predicted with near certainty to be either a noun or a verb, depending on the condition. Four sentence conditions were included: two noun-terminal (“Noun1” and “Noun2” conditions) and two verb-terminal (“Verb1” and “Verb2” conditions), each corresponding to a different structural template as in (1)–(4) below. “Wh” indicates a *wh*-word, “V<sub>aux</sub>” indicates an auxiliary verb (either *did* or *was*), “NP” indicates a noun phrase, “VP” indicates a verb phrase, and “PP” indicates a prepositional phrase.

(1)	Noun1:				
	<i>Wh</i>	V <sub>aux</sub>	NP	PP_____?	
e.g.,	<i>Where</i>	<i>was</i>	<i>the woman</i>	<i>for the _____?</i>	
(2)	Noun2:				
	<i>Wh</i>	V <sub>aux</sub>	NP	VP_____?	
e.g.,	<i>When</i>	<i>did</i>	<i>the janitor</i>	<i>mention the _____?</i>	
(3)	Verb1:				
	<i>Wh</i>	NP	V <sub>aux</sub>	NP	VP_____?
e.g.,	<i>Which</i>	<i>budget</i>	<i>was</i>	<i>the mechanic</i>	<i>permitted to _____?</i>
(4)	Verb2:				
	<i>Wh</i>	NP	V <sub>aux</sub>	NP	VP_____?
e.g.,	<i>What</i>	<i>crib</i>	<i>did</i>	<i>the broker</i>	<i>plan to _____?</i>

In order to control for the possibility that a verb phrase-terminal sentence might accommodate the insertion of a direct object, we constrained verb-terminal sentence completions by using a *wh*-question frame (and in (3) and (4) above). While this necessarily limited the verb completions of verb-terminal sentences to transitive verbs only, the *wh*-type noun phrases such as “*Which budget*” and “*Which crib*” blocked the insertion of a direct object, thus preventing “run-away” phrasal completions. In order to match sentence frames across conditions as much as possible, noun-terminal sentences were also *wh*-questions (see (1) and (2) above).

Noun2-type sentences contained matrix verbs selective for a NP complement ( $\leq 75\%$ ; Jennings, Randall, & Tyler, 1997; Trueswell, Tanenhaus, & Kello, 1993). Matrix verbs of Verb-type sentences were selective for infinitival complements ( $\leq 80\%$ ) and the verbs in Verb1 and Verb2 sentences did not differ significantly in frequency or length (Osterhout & Holcomb, 1992). Sentential subjects for all sentence types were drawn from the same list of agent NPs. We included two different types of sentence frames for both Noun and Verb sentence conditions in order to ensure that any prediction effects would not be specific to any one particular sentence frame. We did not expect any differences between Verb1 and Verb 2 sentences; however, since the Noun1 and Noun2 sentences had a different number of content words, we used this contrast to test whether our ROIs were sensitive to the number of content words in a given sentence (see Results for further discussion).

In order to prevent specific lexical item-based prediction effects from confounding any putative effects of lexical category prediction, sentence stimuli for this study were selected based on the results of a separate web-based sentence completion study that involved a larger set of sentence fragments. Sentence fragments were selected for use in the fMRI study if they had a relatively low Cloze probability but nevertheless still guaranteed either a noun or verb completion. (Cloze probability is the probability that a given sentence frame will end in one particular word.) For the sentence completion study, 37 sentences of each Noun-type condition and 30 sentences of each Verb-type condition were presented, for a total of 134 sentences. (The number of sentences originally generated for each condition was limited by the type of matrix verb used: verbs for the Noun-type sentences were more numerous than for Verb-type sentences.) Seventy-five undergraduates at the University of Pennsylvania, all native speakers of English who received class credit for their participation, were instructed to read the incomplete sentences and type whatever they thought best completed each sentence. Subjects were told to write the first completion that occurred to them, even if that rendered the sentence odd or even nonsensical; otherwise, subjects were not instructed how to limit their answers.

On the basis of the norming study, we removed from the stimuli list any sentence completed with a phrasal constituent of more than one word, and of the sentences that remained, any sentence

completed with the same word item by more than 22 of 75 subjects. Thus the maximum item Cloze probability of the remaining sentences was 29.3%, with the average being 2.8% and 2.9% for noun- and verb-terminal sentences, respectively. Twenty-four sentences in each of the four conditions were then selected for use in the experiment such that no significant differences in CELEX-based frequency or orthographic length existed between the set of noun sentence completions and the set of verb sentence completions (Baayen, Piepenbrock, & Gulikers, 1995).

#### 2.2.2. Noise threshold assessment

Before collecting fMRI data from each subject, we determined the subject-specific level of noise to apply to visual stimuli, using a psychophysical staircase procedure (QUEST staircase technique; Watson & Pelli, 1983) and a customized Matlab script to generate noised images by taking the inverse Fourier transformation of the mean amplitude spectra with randomized phase spectra (Sadr & Sinha, 2004). Images for this procedure consisted of black word tokens on white background. In the staircase session, the subject viewed images the percentage phase coherence of which was gradually reduced until the subject reached threshold on an identification task (80% correct identification, 20 steps by QUEST staircase). This subject-specific thresholded phase coherence was used to generate both the noised word images and the images of pure Gaussian white noise that were presented during fMRI experiment (as in Fig. 1).

#### 2.3. Experimental task and design

The subject’s task on each trial was to read a sentence fragment presented one word at a time centrally and then look for an “appropriate” sentence completion among random dot noise (Fig. 1), where “appropriate” was not specified but left to the subject’s judgment. The subject then would indicate with a button press whether the word that was finally discernible met this criterion. One out of every six sentences presented was ultimately completed ungrammatically (i.e., completed with a noun when a verb was expected, and vice versa), but because even grammatically well-formed sentences were often bizarre (see Appendix, e.g., “What money did the baby start to . . . receive?”), the average judgments for the ostensibly well-formed sentences varied.

In a given trial, between four and eight noise-only images were presented such that the length of time between a cue (the last word of the sentence fragment) and a noise-thresholded target word varied between 12 and 24 s. Subjects were unable to predict the moment the target word would become visible, and therefore had to be vigilant throughout the 12-to-24-s interval. Each word in a given sentence appeared on the screen for 300 ms and was followed by 300 ms of a blank screen ISI. Since Noun-type sentence frames lasted 3600 ms (six words, 300 ms on, 300 ms off) but Verb-type sentence frames lasted 4200 ms (seven words), the



relationship between a word's form and its grammatical category, but that during normal reading, when the syntactic context is predictive of the noun/verb category, form typicality effects are robust. By blocking sentences of similar syntactic category, we aimed to facilitate this sensitivity to predictable syntactic context. Note, however, that in terms of analysis, this was not a block design, but rather a slow event-related design.

#### 2.4. Image acquisition and pre-processing

fMRI data were collected at the Hospital of the University of Pennsylvania on a 3T Siemens Trio System using an eight-channel multiple-array Nova Medical head coil. After acquiring T1-weighted anatomical images (TR = 1620 ms, TE = 3 ms, TI = 950 ms, voxel size = 0.977 mm × 0.977 mm × 1.000 mm), we ran the experimental blocks and collected T2\*-weighted images using a gradient-echo echoplanar pulse sequence (TR = 3000 ms, TE = 30 ms, voxel size = 3 mm × 3 mm × 3 mm). Images were rear-projected onto a Mylar screen at the head of the scanner and viewed through a mirror mounted to the head coil. Words presented on the screen subtended about 5° × 2° of the visual angle and were presented foveally at the center of the screen.

fMRI data were pre-processed offline using the VoxBo ([www.voxbo.org](http://www.voxbo.org)) and AFNI (Cox & Jesmanowicz, 1999) software packages. Voxbo was only used to sort the raw DICOM files for further processing in AFNI. The first four volumes of each functional run were removed so as to allow the signal to reach steady-state magnetization. These initial volumes included a short block in which the target appeared after only 3 or 6 s rather than the full 12–24 s: this encouraged participants to start searching for a target from the very onset of a trial. Functional images were slice-time corrected, and a motion correction algorithm employed in AFNI registered all volumes to a mean functional volume. We applied a high-pass filter of 0.01 Hz on each run to remove low frequency trends. Images were transformed to Talairach standardized space (Talairach & Tournoux, 1988) and voxels were resampled in the process to 3.5 mm × 3.5 mm × 3.5 mm. The data were left unsmoothed for MVPA.

#### 2.5. Analysis

Pattern analyses were implemented in MATLAB using scripts adapted from the Princeton Multi-Voxel Pattern Analysis toolbox (Detre et al., 2006). The analysis pipeline involved three main steps: voxel selection, classifier training, and classifier testing. The latter two steps in the analysis were performed in a 4-fold cross-validation procedure whereby the classifier was trained on three runs and then tested on a fourth in four separate iterations in “leave-one-out procedure” on each individual subject separately (Friedman, Hastie, & Tibshirani, 2001). Below we describe each step of the procedure in detail.

##### 2.5.1. Feature selection

Voxels were selected from several regions of interest (ROIs). The VT mask, which included the fusiform gyrus but also extended to lingual, parahippocampal, and inferior temporal gyri, was drawn for each subject *a la* Haxby (2001). The mid left FG mask was drawn for each subject based on previous reports of visual word form area (VWFA) localization in left mid fusiform gyrus (mid FG) (Cohen & Dehaene, 2004; Cohen et al., 2000, 2002; Hasson, Levy, Behrmann, Hendler, & Malach, 2002; McCandliss, Cohen, & Dehaene, 2003). We used the BA 17 label provided in AFNI's Talairach daemon database (Lancaster et al., 2000) to define the early visual cortex (EVC) mask. Finally, in order to demonstrate the specificity of our findings to visual cortex, we included one additional

ROI, left BA46 (from the same AFNI database), as a negative control region.

We further narrowed these ROIs using the following voxel selection technique. We first created a boxcar regressor for each condition (NOUN and VERB (Noun1 and Noun2 conditions were combined, as were Verb1 and Verb2 conditions)) corresponding to the time points when the subject was predicting, but not actually seeing, a word (separate regressors were also included for sentence presentation, noised-word presentation, and ITI fixation baseline). To preclude the possibility of overlap between bottom-up sensory stimulation of visual areas and top-down activation of these same areas by the preceding sentence or word cue, we only looked at noised images presented at least two TRs (6 s) after the prediction cue (Coutanche & Thompson-Schill, 2014). We then convolved the condition regressors with a hemodynamic response function (gamma-variate) and then computed the per-voxel *F*-statistics for these condition contrasts. Based on this analysis, we selected the *N* voxels (where *N* was 20, 60, 100, or 200) that had the highest *F*-statistics in a given ROI (McDuff, Frankel, & Norman, 2009; Polyn, Natu, Cohen, & Norman, 2005). This procedure was performed for each of the four iterations and only on the three training runs: the test run was always left out in order to avoid “peeking.” Since each iteration chose a different set of *N* voxels, each classifier had a different *N*-unit input layer.

In Fig. 3, we plot classifier performance against each voxel input size (20, 60, 100, 200, where these voxel input sizes were arbitrarily chosen). We report permutation tests for only the 20-voxel inputs.)

##### 2.5.2. Classifier training and testing

The classification analyses reported here used a two-layer neural network classifier as implemented in the Princeton MVPA Matlab toolbox, which itself also used the proprietary Matlab Neural Networks toolbox. Before running the classifier, we z-scored the functional data for each voxel and for each run. We then implemented a simple neural network with an input layer of, e.g., 20 units corresponding to the best 20 voxels (z-scored raw BOLD signal, not GLM beta values), and an output layer, which had two units (for Noun vs. Verb). This was trained on three runs using a conjugate gradient descent backpropagation algorithm (Polyn et al., 2005), and tested on a fourth run in a leave-one-out 4-fold cross-validation procedure.

Classifier weights were initialized at random values and then trained on the three training runs. The order of the training patterns was randomized, and training was stopped when either the output layer's mean cross-entropy error fell to 0.001 or the network made 200 passes through all the training patterns. To average out the variability in classifier performance associated with the random initialization settings, we repeated each fold's classifier procedure 100 times and averaged these output values to yield a single classifier performance output for each fold (McDuff et al., 2009). The folds' performances were then averaged together for a single classifier performance value for each subject.

##### 2.5.3. Assessing classifier performance

To assess the significance of the classifier's accuracy percentages for each individual subject for the 20-voxel inputs, we used a non-parametric statistical procedure to determine whether each individual's classification performance was greater than that expected by chance (Gallivan, McLean, Smith, & Culham, 2011; Golland & Fischl, 2003). The condition labels for each subject were scrambled 100 times, and the classifier was trained and tested on each new set of scrambled labels. These labels were scrambled such that each new scrambled set had the same number of trials and conditions per run as the original set. A group *p*-value was calculated by randomly selecting from each subject's 101 possible

classification scores (100 permuted distributions plus the original “real” classification score), generating a population of 1000 mean accuracies based on 1000 combinations of randomly drawn classification scores. The real mean group classification score was then compared to this permuted null distribution of 1000 means to identify the  $p$ -value.

### 3. Results

In a univariate test of differential activity in all ROIs between noun-type and verb-type predictions, no clusters survived a significance threshold of  $p=0.1$ . However, multivariate analyses revealed a difference between noun and verb predictions in a sub-region of VT corresponding to the left mid FG, inclusive of an area sometimes called the “visual word form area.” Average classification performance in the left mid FG mask was significantly above 50% chance ( $M = 0.58$ , two-tailed permutation test  $p = 0.023$ ; 20-voxel input). Group-level classification was highest for a 20-voxel input in mid FG, but classification was also reliably above chance for 60 and 100 voxels in mid FG. In order to examine whether mid FG may be driving classification in VT, we ran a separate classifier over the complement set of voxels that were in VT but not in mid FG (see Fig. 2). This VT complement classification did not reach significance at the group level ( $M = 0.53$ , permutation test  $p = 0.14$ ), consistent with the hypothesis that within the left VT ROI, mid FG is particularly sensitive to word form features underlying the noun–verb distinction. Finally, classification of nouns vs. verbs did not meet significance in either EVC (see Fig. 3;  $M = 0.50$ , permutation test  $p = 0.41$ ) or BA46 ( $M = 0.51$ ,  $p = 0.20$ ). These data

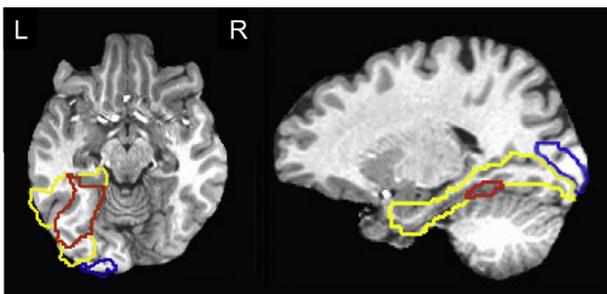
together indicate that the left mid FG is tuned to lexical-syntactic features of words in predictive contexts.

We next considered the possibility that successful classification could have been the result of a confound between sentence type and sentence length (specifically, the number of content words), even though subjects were not viewing the sentences when the relevant fMRI data were being collected. When we created sentences for this study, we decided that maintaining suitable syntactic structure was more desirable than matching the two conditions on length, so we designed the experiment to allow us to test for this possible length confound. Specifically, the first type of noun-terminal sentence (Noun1) had one content word per sentence, the second type of noun-terminal sentence (Noun2) had two content words, and both types of verb-terminal sentences (Verb1 and Verb2) had three content words. In order to determine whether the number of content words could drive classification, we tested classification on the two types of noun-terminal sentences (i.e., same grammatical class, but one content word vs. two content words). Classification was not significant in VT (Noun1 vs. Noun2:  $M = 0.48$ ,  $p = 0.52$ , 20 voxels). Because this null result may simply have been due to insufficient power (since the number of volumes used to train each pattern was half that used in the primary analysis), we also trained classifiers on the Noun1-vs.-Verb1 and the Noun2-vs.-Verb2 comparisons in VT. VT voxels in mid FG were able to train classifiers on Noun1-vs.-Verb1 above chance ( $M = 0.54$ ,  $p < 0.01$ ). This pattern gives us confidence that classification performance for Nouns vs. Verbs was due to differences in lexical-syntactic prediction and not sentence length.

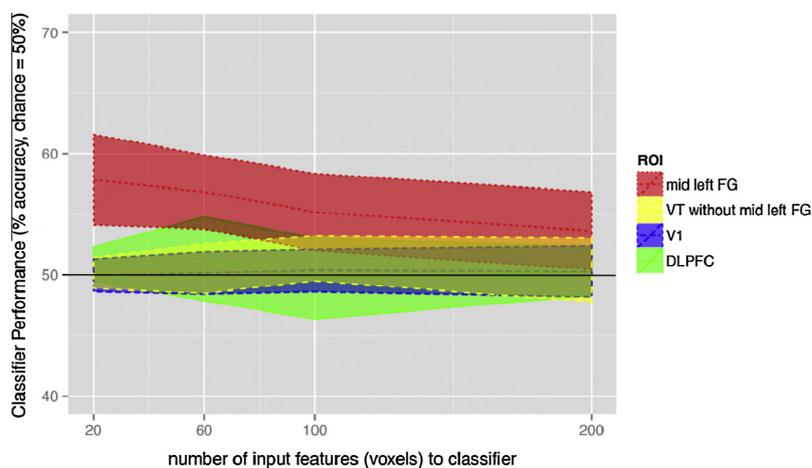
### 4. Discussion

Using multi-voxel pattern analysis (MVPA), we were able to show that participants’ expectations of a noun or a verb can generate form-based predictions in higher level visual areas, including the left ventral temporal (VT) cortex, and a subregion of the latter – the left mid fusiform gyrus (mid FG), sometimes called the “visual word form area.” No evidence for form-based predictions was observed in early visual cortex (EVC).

What differences between nouns and verbs might give rise to their successful classification in VT and mid FG? Nouns and verbs differ along a number of dimensions – phonological, lexical, syntactic, and semantic – but because left mid FG was the primary locus of noun–verb classification in VT and has been shown to be most sensitive to bigrams and lexical-level properties of words, it is likely that word form features are indeed the representational



**Fig. 2.** ROI outlines. Extent of EVC (BA 17) shown in blue, VT in yellow, and mid FG in red in representative subject. Note mid FG is a subregion of VT and thus the borders of these ROIs overlap. Axial slice shown at  $z = -10$ , sagittal slice shown at  $x = -18$ .



**Fig. 3.** Classification across subjects ( $n = 10$ ) over all folds ( $n = 4$ ) using best 20, 60, 100, and 200 – voxels per fold. Lines indicate median classification performance, and ribbons indicate confidence intervals of classification performance across the ten subjects.

substrate of noun–verb classification. While it has been found that nouns are more highly imageable than verbs (e.g., Luzzatti & Chierchia, 2002), effects of noun and verb imageability are more often found in the right hemisphere (Crepaldi et al., 2006). Furthermore, in testing for gross differences between nouns and verbs such as imageability, our univariate analysis of noun prediction vs. verb prediction in both left mid FG and VT found no significant differential activation between these two grammatical categories. If imageability, or some other coarse visual property of nouns and verbs deriving from their conceptual semantic associations, were underlying classification, then we would expect to see this borne out in the univariate test. Thus, it is very likely that we are classifying word effects at a finer grain.

#### 4.1. Syntactic prediction, or something else?

The nature of the prediction that generates the form-based estimate in these visual areas is far from clear. Indeed, we cannot preclude the possibility that visual word form estimates arise not from syntactic cues, but via more general conceptual semantic properties of nouns and verbs intervening between a syntactic category prediction and the word form estimation. It is also possible that neither syntactic nor semantic information is necessary to make the sort of noun-vs.-verb predictions seen in this study. Note that the penultimate word in all Noun sentences was “the,” while the penultimate word in all “Verb” sentences was “to.” To ensure we did not sample bottom-up signal changes associated with processing the word forms “to” or “the,” we did not include the first two TRs after the prediction cue. Thus the single word forms “to” and “the” are unlikely to drive classification at the sensory-perceptual level. However, that is not the only way “to/the” might have driven classification. There is also the possibility that the subject only needed the “to/the” cue, not the sentence frame, to generate a prediction. That is, it may be possible to generate a noun- or verb-prediction on the basis of the bigrams associated with “to” and “the,” without recourse to a syntactic structure per se. This is unlikely, however, since the subject could not complete the task using only the single word cue: the task was an acceptability judgment, not a grammaticality judgment, and some grammatical sentences were deemed unacceptable by most subjects (e.g., “Who was the boss from the ... lawyer?”) Similarly, it is unlikely we are classifying a rehearsal effect of “to” vs. “the,” since this would not be sufficient for task completion.

We have further reason to believe syntactic structure, and not lower-level prediction driven by single-word prediction, accounts for these noun-vs.-verb prediction results. In a separate study, we presented subjects with specific noun-typical nouns and verb-typical verbs to search for in noisy dot patterns. This study was identical in design and format to the one reported here, but the cue was a specific word (e.g., “movie”) rather than a sentence fragment. We were unable to classify nouns vs. verbs above chance in VT or mid FG. Despite the greater precision involved in predicting a specific word, classification of sentence-cued prediction outperformed word-cued prediction of nouns vs. verbs. While the brain is likely generating predictions at multiple levels – semantic, syntactic, phonological, orthographic – the presence of sentence structure may direct attention away from irrelevant form-feature information and toward those form-features most diagnostic of nouns and verbs; thus, syntactic structure might serve to tune word form prediction.

#### 4.2. Implications for models of syntactic prediction and word form feature estimation

The speed at which the human brain processes language is remarkable, and it is becoming increasingly clear that models of

strictly serialized, modular, bottom-up language processing are insufficient to account for this performance. Language operates on a number of representational levels, such modules including phonology, syntax, semantics, etc., but it is not clear how possible interstratal relationships among these levels may affect processing. One such interstratal relationship – that between the phonology (read out as orthography) and lexical-syntactic category of a word – has been largely overlooked. The possibility that lexical-syntactic categories have form-feature signatures at much lower levels of representation may obscure our measure of when and where syntactic effects occur. For instance, many neuro-cognitive accounts of syntactic processing derive from electrophysiological studies reporting that certain syntactic factors affect processing in an earlier time window than lexical-semantic violations. Event-related potentials (ERPs) modulated by syntactic violations have been observed as early as ~60 ms after the onset of an unexpected word (many studies find a so-called Early Left-Anterior Negativity at ~125 ms post stimulus onset (Friederici, Pfeifer, & Hahne, 1993; Neville, Nicol, Barss, Forster, & Garrett, 1991)), while others find a mismatch negativity even prior to this at ~60 ms p.s.o. (Herrmann, Maess, Hahne, Schröger, & Friederici, 2011; Herrmann, Maess, Hasting, & Friederici, 2009). In contrast, neural correlates of lexical, semantic, and world knowledge violations are typically indexed by a negative deflection of the ERP much later at around 400 ms post-stimulus onset (the N400 component (Kutas & Hillyard, 1980)). Many researchers interpret this pattern as evidence of a “syntax-first” model of language comprehension, where the initial stage of processing reflects syntactic computation, only after which lexical and semantic factors are accessed.

Friederici’s (2002) model, for instance, portrays different levels of analysis occurring in an explicitly serial, modular, and bottom-up fashion, and that identification of lexical-syntactic category occurs at about 150–200 ms (Friederici, 2002). However, our study, among others, now raises the possibility that some of these putatively syntactic effects may fall out of form typicality effects, modulated by top-down syntactic prediction. While the prediction effect we see may in some sense be syntactic, it certainly does not entail bottom-up syntactic analysis of the word being predicted, but rather a form-feature prediction based on the syntactic structure already built. Therefore, in order to distinguish syntactic analysis per se from downstream effects of syntactic prediction, future studies investigating early syntactic effects should account for the form-feature properties of words in addition to manipulating syntactic context.

#### 4.3. Neural circuitry of visual prediction

There is a growing body of research investigating top-down effects on sensory processing, but most evidence for such expectation-induced processing comes from studies showing pre-activation of brain regions subserving rather coarse domains of sensory information; e.g., gustatory cortex activating when subjects expect food items (Simmons, Martin, & Barsalou, 2005), somatosensory cortex activating during subjects’ anticipation of somatosensory stimuli (Carlsson, Petrovic, Skare, Petersson, & Ingvar, 2000), and fusiform face area activating when subjects expect faces, as opposed to other objects (Summerfield et al., 2006). Likewise, when subjects are anticipating some reading task, we would expect to see enhanced activation of visual areas as opposed to, say, auditory or somatosensory regions. However, while these sorts of study provide interesting fodder for models of attention, they do not themselves offer evidence for predictive coding of particular percepts within these sensory domains, where “predictive coding” refers to a model whereby prior information facilitates top-down conditional expectations at the level of the sensorium (Friston, 2003). That is, just because visual attention may be upregulated

during reading does not mean that visual cortex has any predictions about the upcoming word.

However, studies manipulating attentional demands during predictive vs. non-predictive tasks have shed light on the neural connectivity underlying expectation-based visual processing. For instance, Summerfield and Koechlin (2008) found increased backward connectivity from FG to EVC in prediction conditions of low-level visual (non-linguistic) stimuli, but not non-predictive conditions. Conversely, they found increased forward connectivity from EVC to FG in cases of prediction mismatch, suggesting FG is a hub for both generating and updating visual form-based predictions. These data provided evidence for attentional modulation of FG depending on whether the task was predictive or not, but did not offer a means to detect whether FG was engaged in predictive coding per se.

In the current study, we probed both FG and EVC for predictive coding of nouns vs. verbs, but only found significant classification in FG (and only when prediction was based on syntactic structure). This is consistent with an account whereby FG mediates prediction of visual features when prediction is maximally afforded. Failure to classify nouns and verbs in EVC would be wholly unsurprising were it not for previous evidence from MEG that early occipital activity is sensitive to nouns' form typicality (Dikker et al., 2010). However, there are two caveats regarding the MEG evidence. First, the source of the ERF was estimated using equivalent current dipole analysis, a localization technique requiring an assumption on the part of the investigator as to the number of sources contributing to a given field pattern. Second, while this technique is generally suitable for analyzing early sensory components such as the M100, it is possible that both the waveform and the scalp topography of the relevant magnetic field was distorted by use of a relatively high cut-off of 1 Hz when high-pass filtering (Acunzo, MacKenzie, & van Rossum, 2012). Thus, while the MEG data provide evidence of early sensitivity to noun form typicality, the effect cannot be unequivocally localized to EVC.

## 5. Conclusions

The phenomenon of form typicality currently provides the best account of how syntactic categories can map onto word forms. However, it may be that the sentence-context cues in this study allow for better prediction of nouns vs. verbs because they sustain attention better – or longer – than single-word cues (see Section 4.1). That sentence context appears to be privileged may be an accident of the temporal structure of the sentence cue rather than a function of the syntactic representation itself, and the current study cannot differentiate between these two possibilities.

It should also be noted that the current study cannot distinguish between a model involving true predictive coding (i.e., where a syntactic prediction is transformed to a phonological/orthographic form feature distribution) and a model whereby attention to a certain subset of noun-diagnostic features is upregulated when a noun is expected (vs. a verb), and vice versa for verb prediction. The difference between these two models is subtle, and computationally these two models may be indistinguishable: both entail pre-activation of certain features based on syntactic information, and both entail lateral inhibition of irrelevant features. Rather, the distinction lies in our understanding of noun/verb form typicality, and it may be that the notion of “form typicality” as currently delineated encompasses more features than are strictly necessary for distinguishing noun and verb forms. A more explicit definition of phonological/orthographic “form typicality,” at least insofar as it discriminates nouns from verbs, is a crucial desideratum for a future theory of word form prediction.

## Acknowledgments

Thanks to the labs of Sharon Thompson-Schill and John True-swell for their advice and input. Special thanks to Marc Coutanche for aiding in design and analyses. This research was supported by grants R01MH70850, R01EY021717 (Sharon Thompson-Schill) and NNC P30 NS45839 (Center for Functional Neuroimaging).

## Appendix A

Sentence fragment stimuli with final completion word, which was presented in noise.

1. \*Who was the broker from the ... include?
2. Where was the doctor for the ... baby?
3. Who was the reporter in the ... story?
4. Who was the senator from the ... movie?
5. Who was the singer in the ... band?
6. Where was the teacher with the ... child?
7. Which crib did the broker plan to ... destroy?
8. \*Which opera did the reporter agree to ... sofa?
9. What sermon did the woman struggle to ... believe?
10. Which test did the senator agree to ... take?
11. What band did the judge hope to ... accuse?
12. Which army did the tailor hope to ... assist?
13. \*What bank was the doctor implored to ... bible?
14. What child was the broker persuaded to ... adopt?
15. What money was the reporter selected to ... spend?
16. Which patient was the senator forced to ... abduct?
17. What story was the judge advised to ... review?
18. Which course was the tailor hired to ... take?
19. When did the broker accept the ... money?
20. \*Where did the woman forget the ... amuse?
21. When did the senator learn the ... news?
22. Where did the judge maintain the ... order?
23. When did the tailor observe the ... murder?
24. Where did the general recall the ... story?
25. \*Where did the mechanic teach the ... lend?
26. When did the minister reveal the ... bible?
27. Where did the teacher demand the ... salary?
28. When did the burglar discover the ... house?
29. Where did the dentist stress the ... teeth?
30. When did the janitor mention the ... flood?
31. Which budget was the mechanic permitted to ... accept?
32. What turtle was the teacher urged to ... adopt?
33. Which student was the burglar ordered to ... remove?
34. What product was the policeman persuaded to ... examine?
35. \*Which defendant was the dentist invited to ... marble?
36. What committee was the janitor bribed to ... attend?
37. \*What car did the general refuse to ... bible?
38. Which bird did the minister start to ... adopt?
39. What patient did the mechanic refuse to ... assist?
40. Which paper did the singer decide to ... read?
41. What clinic did the teacher plan to ... attend?
42. What story did the dentist hope to ... write?
43. Who was the burglar from the ... movie?
44. Where was the dentist for the ... lawyer?
45. Where was the nurse with the ... child?
46. Where was the journalist from the ... paper?
47. \*Where was the governor for the ... lend?
48. Who was the nephew in the ... movie?
49. What staircase was the prince encouraged to ... use?
50. Which book was the man hired to ... read?
51. What store was the governor convinced to ... close?
52. Which soldier was the nephew persuaded to ... respect?

53. What doctor was the salesman induced to ... hire?
54. Which country was the executive ordered to ... bomb?
55. Where did the nurse accept the ... money?
56. When did the journalist advise the ... editor?
57. \*Where did the governor confirm the ... include?
58. When did the nephew forget the ... rumor?
59. \*Where did the salesman learn the ... adopt?
60. When did the executive maintain the ... order?
61. Where was the salesman with the ... bible?
62. Who was the executive from the ... firm?
63. Who was the writer from in the ... movie?
64. Where was the woman with the ... baby?
65. Who was the mother from the ... movie?
66. Where was the baby for the ... diaper?
67. Which company did the janitor try to ... join?
68. What lesson did the governor decide to ... give?
69. \*What play did the salesman try to ... marble?
70. Which child did the executive plan to ... adopt?
71. What transaction did the professor try to ... amuse?
72. Which faucet did the writer decide to ... use?
73. \*What actress did the woman struggle to ... movie?
74. Which mayor did the mother agree to ... marry?
75. What money did the baby start to ... receive?
76. \*Which job did the lawyer want to ... sofa?
77. What drug did the boss want to ... abuse?
78. What insect did the juggler desire to ... eat?
79. Who was the man in the ... line?
80. Who was the prince from the ... castle?
81. Who was the boss from the ... lawyer?
82. Where was the banker for the ... deal?
83. Where was the aunt with the ... money?
84. Who was the nanny from the ... story?
85. \*When did the lawyer demand the ... include?
86. Where did the boss discover the ... drugs?
87. When did the banker stress the ... truth?
88. \*Where did the juggler mention the ... adopt?
89. When did the man accept the ... cash?
90. Where did the policeman advise the ... lawyer?
91. What palace was the professor allowed to ... inhabit?
92. Which engine was the writer urged to ... avoid?
93. Which vase was the baby permitted to ... break?
94. Which bird was the nanny scared to ... touch?
95. What statue was the gymnast required to ... avoid?
96. Which umbrella was the nun meant to ... carry?

## References

- Acunzo, D. J., MacKenzie, G., & van Rossum, M. C. W. (2012). Systematic biases in early ERP and ERF components as a result of high-pass filtering. *Journal of Neuroscience Methods*, *201*, 66–81. <http://dx.doi.org/10.1016/j.jneumeth.2012.06.011>.
- Arciuli, J., & Monaghan, P. (2009). Probabilistic cues to grammatical category in English orthography and their influence during reading. *Scientific Studies of Reading*, *13*, 73–93.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database*. Philadelphia: Linguistic Data Consortium.
- Carlsson, K., Petrovic, P., Skare, S., Petersson, K. M., & Ingvar, M. (2000). Ticking expectations: Neural processing in anticipation of a sensory stimulus. *Journal of Cognitive Neuroscience*, *12*(4), 691–703. <http://www.mitpressjournals.org/doi/abs/10.1162/0899892900562318>.
- Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, *30*(3), 348–369. <http://www.sciencedirect.com/science/article/pii/0749596X9190041H>.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
- Cohen, L., & Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. *NeuroImage*, *22*(1), 466–476. <http://www.sciencedirect.com/science/article/pii/S1053811904000576>.
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A., et al. (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*(2), 291–307. <http://dx.doi.org/10.1093/brain/123.2.291>.
- Cohen, L., Lehéricy, S., Chochon, F., Lemer, C., Rivaud, S., & Dehaene, S. (2002). Language-specific tuning of visual cortex? Functional properties of the visual word form area. *Brain*, *125*(5), 1054–1069. <http://dx.doi.org/10.1093/brain/awf094>.
- Coutanche, M. N., & Thompson-Schill, S. L. (2014). Creating concepts from converging features in human cortex. *Cerebral Cortex*. <http://dx.doi.org/10.1093/cercor/bhu057>.
- Cox, R. W., & Jesmanowicz, A. (1999). Real-time 3D image registration for functional MRI. *Magnetic Resonance in Medicine*, *42*(6), 1014–1018. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.43.253&rep=rep1&type=pdf>.
- Crepaldi, D., Aggujaro, S., Arduino, L. S., Zonca, G., Ghirardi, G., Inzaghi, M. G., et al. (2006). Noun–verb dissociation in aphasia: The role of imageability and functional locus of the lesion. *Neuropsychologia*, *44*(1), 73–89. <http://dx.doi.org/10.1016/j.neuropsychologia.2005.04.006>.
- De Saussure, F. (1916). *Cours de Linguistique Generale*. Paris: Payot.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, *15*(6), 254–262.
- Detre, G. J., Polyn, S. M., Moore, C. D., Natu, V. S., Singer, B. D., Cohen, J. D., et al. (2006). The Multi-Voxel Pattern Analysis (MVPA) toolbox. *Human Brain Mapping*.
- Dikker, S., Rabagliati, H., Farmer, T. A., & Pyllkanen, L. (2010). Early occipital sensitivity to syntactic category is based on form typicality. *Psychological Science*, *21*(5), 629–634. <http://dx.doi.org/10.1177/0956797610367751>.
- Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences*, *103*(32), 12203–12208.
- Farmer, T. A., Monaghan, P., Misyak, J. B., & Christiansen, M. H. (2011). Phonological typicality influences sentence processing in predictive contexts: Reply to Staub, Grant, Clifton, and Rayner (2009). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1318–1325. <http://dx.doi.org/10.1037/a0023063>.
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78–84.
- Friederici, A. D., Pfeifer, E., & Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research*, *1*(3), 183–192. <http://www.sciencedirect.com/science/article/pii/0926641093900262>.
- Friedman, J., Hastie, T., & Tibshirani, R. (2001). *The elements of statistical learning* (Vol. 1). Springer Series in Statistics. <http://www-stat.stanford.edu/~tibs/book/preface.ps>.
- Friston, K. (2003). Learning and inference in the brain. *Neural Networks*, *16*(9), 1325–1352. <http://dx.doi.org/10.1016/j.neunet.2003.06.005>.
- Gallivan, J. P., McLean, D. A., Smith, F. W., & Culham, J. C. (2011). Decoding effector-dependent and effector-independent movement intentions from human parieto-frontal brain activity. *Journal of Neuroscience*, *31*(47), 17149–17168. <http://dx.doi.org/10.1523/JNEUROSCI.1058-11.2011>.
- Golland, P., & Fischl, B. (2003). Permutation tests for classification: Towards statistical significance in image-based studies. *Information Processing in Medical Imaging*, *18*, 330–341.
- Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, *34*(3), 479–490. <http://www.sciencedirect.com/science/article/pii/S0896627302006621>.
- Haushofer, J., Livingstone, M. S., & Kanwisher, N. (2008). Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. *PLoS Biology*, *6*(7), e187. <http://dx.doi.org/10.1371/journal.pbio.0060187>.
- Haxby, J. V. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*(5539), 2425–2430. <http://dx.doi.org/10.1126/science.1063736>.
- Herrmann, B., Maess, B., Hahne, A., Schröger, E., & Friederici, A. D. (2011). Syntactic and auditory spatial processing in the human temporal cortex: An MEG study. *NeuroImage*, *57*(2), 624–633. <http://dx.doi.org/10.1016/j.neuroimage.2011.04.034>.
- Herrmann, B., Maess, B., Hasting, A. S., & Friederici, A. D. (2009). Localization of the syntactic mismatch negativity in the temporal cortex: An MEG study. *NeuroImage*, *48*(3), 590–600. <http://dx.doi.org/10.1016/j.neuroimage.2009.06.082>.
- Hunt, R. H., & Aslin, R. N. (2010). Category induction via distributional analysis: Evidence from a serial reaction time task. *Journal of Memory and Language*, *62*(2), 98–112.
- Itier, R. J., Herdman, A. T., George, N., Cheyne, D., & Taylor, M. J. (2006). Inversion and contrast-reversal effects on face processing assessed by MEG. *Brain Research*, *1115*(1), 108–120. <http://dx.doi.org/10.1016/j.brainres.2006.07.072>.
- Jennings, F., Randall, B., & Tyler, L. K. (1997). Graded effects of verb subcategory preferences on parsing: Support for constraint-satisfaction models. *Language and Cognitive Processes*, *12*(4), 485–504. Retrieved from: <http://www.tandfonline.com/doi/abs/10.1080/016909697386817>.
- Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, *99*(2), 349. <http://psycnet.apa.org/journals/rev/99/2/349>.
- Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*(4427), 203–205. <http://dx.doi.org/10.2307/1683915>.

- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., et al. (2000). Automated Talairach Atlas labels for functional brain mapping. *Human Brain Mapping, 10*(3), 120–131. [http://dx.doi.org/10.1002/1097-0193\(200007\)10:3<120::AID-HBM30>3.0.CO;2-8](http://dx.doi.org/10.1002/1097-0193(200007)10:3<120::AID-HBM30>3.0.CO;2-8).
- Lee, S.-H., Kravitz, D. J., & Baker, C. I. (2012). Disentangling visual imagery and perception of real-world objects. *NeuroImage, 59*(4), 4064–4073. <http://dx.doi.org/10.1016/j.neuroimage.2011.10.055>.
- Luzzatti, C., & Chierchia, G. (2002). On the nature of selective deficits involving nouns and verbs. *Italian Journal of Linguistics, 14*, 43–72. <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.175.2377&rep=rep1&type=pdf>>.
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7*(7), 293–299. [http://dx.doi.org/10.1016/S1364-6613\(03\)00134-7](http://dx.doi.org/10.1016/S1364-6613(03)00134-7).
- McDuff, S. G. R., Frankel, H. C., & Norman, K. A. (2009). Multivoxel pattern analysis reveals increased memory targeting and reduced use of retrieved details during single-agenda source monitoring. *The Journal of Neuroscience, 29*(2), 508–516.
- Monaghan, P., Christiansen, M., & Chater, N. (2007). The phonological–distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology, 55*(4), 259–305. <http://dx.doi.org/10.1016/j.cogpsych.2006.12.001>.
- Neville, H., Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*(2), 151–165. <http://dx.doi.org/10.1162/jocn.1991.3.2.151>.
- Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*(6), 785–806. Retrieved from: <<http://www.sciencedirect.com/science/article/pii/0749596X9290039Z>>.
- Peelen, M. V., & Kastner, S. (2011). A neural basis for real-world visual search in human occipitotemporal cortex. *Proceedings of the National Academy of Sciences, 108*(29), 12125–12130. <http://dx.doi.org/10.1073/pnas.1101042108>.
- Polyn, S. M., Natu, V. S., Cohen, J. D., & Norman, K. A. (2005). Category-specific cortical activity precedes retrieval during memory search. *Science, 310*(5756), 1963–1966. <http://dx.doi.org/10.1126/science.1117645>.
- Sadr, J., & Sinha, P. (2004). Object recognition and random image structure evolution. *Cognitive Science, 28*(2), 259–287. <http://dx.doi.org/10.1016/j.cogsci.2003.09.003>.
- Simmons, W. K., Martin, A., & Barsalou, L. W. (2005). Pictures of appetizing foods activate gustatory cortices for taste and reward. *Cerebral Cortex, 15*(10), 1602–1608. <http://dx.doi.org/10.1093/cercor/bhi038>.
- Solomyak, O., & Marantz, A. (2009). Lexical access in early stages of visual word processing: A single-trial correlational MEG study of heteronym recognition. *Brain and Language, 108*(3), 191–196. <http://dx.doi.org/10.1016/j.bandl.2008.09.004>.
- Staub, A., Grant, M., Clifton, C., & Rayner, K. (2009). Phonological typicality does not influence fixation durations in normal reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(3), 806–814. <http://dx.doi.org/10.1037/a0015123>.
- Stokes, M., Thompson, R., Cusack, R., & Duncan, J. (2009). Top-down activation of shape-specific population codes in visual cortex during mental imagery. *Journal of Neuroscience, 29*(5), 1565–1572. <http://dx.doi.org/10.1523/JNEUROSCI.4657-08.2009>.
- Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 106*(46), 19569–19574. <http://dx.doi.org/10.1073/pnas.0905306106>.
- Summerfield, C., Egnér, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science, 314*(5803), 1311–1314. <http://dx.doi.org/10.1126/science.1132028>.
- Summerfield, C., & Koechlin, E. (2008). A neural representation of prior information during perceptual inference. *Neuron, 59*(2), 336–347. <http://dx.doi.org/10.1016/j.neuron.2008.05.021>.
- Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. New York: Thieme Medical Publishers Inc.
- Tanenhaus, M. K., & Hare, M. (2007). Phonological typicality and sentence processing. *Trends in Cognitive Sciences, 11*(3), 93–95.
- Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain, 122*(11), 2119–2132. <http://dx.doi.org/10.1093/brain/122.11.2119>.
- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*(3), 528–553. <http://dx.doi.org/10.1037/0278-7393.19.3.528>.
- Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., 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*(1), 143–156. <http://dx.doi.org/10.1016/j.neuron.2007.05.031>.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Attention, Perception, & Psychophysics, 33*(2), 113–120. <<http://www.springerlink.com/index/L8V5131R21Q3645W.pdf>>.