Syntactic processing in the human brain: What we know, what we don't know, and a suggestion for how to proceed

Published in final edited form as:
Brain Lang. 2012 February ; 120(2): 187–207. doi:10.1016/j.bandl.2011.01.001.

Syntactic processing in the human brain: What we know, what we don't know, and a suggestion for how to proceed

Evelina Fedorenko, **Alfonso Nieto-Castañón**, and **Nancy Kanwisher**
McGovern Institute for Brain Research, MIT

Abstract

For every claim in the neuroimaging literature about a particular brain region supporting syntactic processing, there exist other claims implicating the target region in different linguistic processes, and, in many cases, in non-linguistic cognitive processes (e.g., Blumstein, 2009). We argue that traditional group analysis methods in neuroimaging may obscure functional specificity because of inter-subject anatomical variability (Fedorenko & Kanwisher, 2009). In Fedorenko et al. (2010) we presented a functional localizer that allows quick and reliable identification of key language-sensitive regions in each individual brain. This approach enables pooling data from corresponding functional regions across subjects rather than from the same locations in stereotaxic space that may differ functionally due to inter-subject anatomical variability. In the current paper we demonstrate that the individual-subjects functional localization approach is superior to the traditional methods in its ability to distinguish among conditions in a brain region's response. This ability is at the core of all neuroimaging research and is critical for answering questions of functional specialization (e.g., does a brain region specialize for processing syntactic aspects of the linguistic signal), which is in turn essential for making inferences about the precise computations conducted in each brain region. Based on our results, we argue that supplementing existing methods with an individual-subjects functional localization approach may lead to a clearer picture of the neural basis of syntactic processing, as it has in some other domains, such as high-level vision (e.g., Kanwisher, 2010) and social cognition (e.g., Saxe & Kanwisher, 2003).

Introduction

Humans share with other species the ability to learn a mapping between strings of sounds and meanings (e.g., Savage-Rumbaugh et al., 1993; Pepperberg, 2000; Kaminsky et al., 2004). However, only humans can (a) combine these lexical-level representations in novel ways in order to convey how objects relate to one another in the world or to describe a state of one's own or another's mind, or (b) infer the relationships among elements in a string of words (cf. Gentner et al., 2006; Outtara et al., 2009). Because this ability – syntactic processing – is one of the defining characteristics of our species, scientists have long sought to understand the cognitive and neural architecture of the system underlying it, as well as its relationship to other human cognitive abilities. In the current paper, we focus on the neural machinery supporting syntactic processing, with a particular focus on investigations using functional magnetic resonance imaging (fMRI), and we argue that a new method – defining regions of interest functionally in individual subjects and then rigorously investigating the

Send correspondence to: Evelina Fedorenko, MIT 46-4141C, Cambridge, MA 02139, USA, evelina9@mit.edu.

functional properties of these regions – may be promising for future investigations of this and related questions.

We begin with a brief summary of what is currently known about the neural basis of syntactic processing from neuroimaging studies, and we outline some of the questions that remain unresolved. We point out that all neuroimaging is premised on the ability to detect in the relevant brain region a differential response to conditions that differ in the cognitive operations they entail, so maximizing this ability (referred to here as "functional selectivity") is of paramount importance in all functional imaging studies. We argue that traditional group analysis methods based on voxels defined in a common stereotaxic space may obscure functional selectivity because subjects differ from each other anatomically. We further argue that an individual-subjects functional localization approach presents a powerful complement to these traditional methods, and may lead to a clearer picture of the neural basis of syntactic processing, as it has in some other domains, such as high-level vision (e.g., Kanwisher, 2010) and social cognition (e.g., Saxe & Kanwisher, 2003). In support of this position, we directly compare the individual-subjects functional localization approach to the traditional approaches with respect to their functional selectivity, using a dataset from Fedorenko et al. (2010). We focus on regions that have been argued to be engaged in syntactic processing based on previous studies. Based on the higher selectivity observed in functionally defined subject-specific regions of interest (ROIs) compared to several kinds of standard group ROIs1, we conclude that functional localization in individual subjects holds promise for understanding the neural basis of language.

The neural basis of syntactic processing: What we know and what we don't know

Several brain regions have been consistently implicated in structural processing. We here briefly review the main candidates (see e.g., Kaan & Swaab, 2002; Friederici & Kotz, 2003; Grodzinsky & Friederici, 2006; Bornkessel-Schlesewsky & Friederici, 2007; Caplan, 2007, for reviews).

In a seminal finding, Caramazza & Zurif (1976) demonstrated that Broca's agrammatic aphasics have difficulty interpreting structures where the order of the noun phrases does not correspond to the typical order of thematic roles in English (for example, reversible passives like *The girl was chased by the boy*, where the sentence-initial noun phrase *the girl* occupies the position of a typical agent but is assigned the patient role by the verb phrase). Ever since this discovery, Broca's area has been a focus of investigations of syntactic processing. The traditional definition of this region includes opercular and triangular portions of the left inferior frontal gyrus, IFG (cf. Lindenberg et al., 2007, for an extensive discussion of the variability in what researchers refer to as "Broca's area" across neuroimaging studies). Regions in/around the left IFG have indeed been shown to respond more strongly to syntactically complex sentences across a number of studies (e.g., Stromswold et al., 1996; Ben-Shachar et al., 2003, 2004; Grewe et al., 2005; Friederici et al., 2006; Santi & Grodzinsky, 2007). However, other studies have implicated these regions in lexico-semantic processing (e.g., Hagoort et al., 2004, 2009; Rodd et al., 2005; Schnur et al., 2009) and phonological processing (e.g., Blumstein et al., 2005; Myers et al., 2009). Some have claimed that different portions of the left IFG support different aspects of linguistic processing (e.g., Buckner et al., 1995; Fiez, 1997; Poldrack et al., 1999; Bokde et al., 2001; Burton, 2001, 2009; Dapretto & Bookheimer, 1999; Devlin et al., 2003; Gough et al., 2005; Costafreda et al., 2006; Santi & Grodzinsky, 2007; Heim et al., 2009), or even different aspects of syntactic processing (e.g., Friederici et al., 2006), but no consensus has yet been

[1]Throughout the paper, we use the term "standard group ROIs" to refer to cases where the same set of voxels in stereotaxic space is used across subjects.

reached about what the relevant divisions are and what cognitive function(s) each sub-region supports (cf. alternative accounts that postulate broader cognitive functions to account for a range of findings; e.g., Novick et al., 2005; Duncan, 2001, 2010; Badre & D'Esposito, 2009).

Posterior temporal regions (posterior superior temporal gyrus/sulcus, STG/STS) have also been implicated in syntactic processing (e.g., Schlesewsky & Bornkessel, 2004; Cooke et al., 2002; Constable et al., 2004; Hasson et al., 2006). Grodzinsky and Friederici (2006) in a recent review hypothesize that these regions may be critical for integrating incoming elements into the evolving structure. However, other studies have observed activations in these regions for non-syntactic language tasks (e.g., Noppeney et al., 2004; Graves et al., 2008; Hickok et al., 2009).

Another region that has been reported in a few studies to be sensitive to syntactic processing is located in the left anterior temporal lobe (ATL). For example, Noppeney & Price (2004) observed a decreased response in this region when participants read sentences with similar structures (see also Friederici et al., 2000a, 2003; Humphries et al., 2005, 2006; Rogalsky & Hickok, 2009, among others). However, as with the other regions, syntactic processing is not the only cognitive function that has been observed to activate ATL structures. For example, Patterson et al. (2007) have argued that these structures support a modal conceptual representations, although some (e.g., Humphreys et al., 2001) have argued that these regions are more engaged during sentence understanding than non-linguistic conceptual processing.

In summary, each of the regions that have been implicated in syntactic processing (including the regions mentioned above as well as some other ones) has also been implicated in other linguistic functions. Moreover, most of these regions have also been implicated in a variety of non-linguistic functions (e.g., see Duncan, 2001, 2010; Lindenberg, 2007; Badre & D'Esposito, 2009; Fadiga et al., 2009, for reviews of tasks that activate lateral frontal cortices; see Hein & Knight, 2008, for a review of tasks that activate STS/STG regions; see Olson et al., 2007, for a review of findings on the role of left temporal pole structures in social/emotional processing). This situation has recently been synopsized in a review by Blumstein (2009), as follows: "The functional properties of language, that is, speech, lexical processing, and syntactic processing, appear not to be focally represented in one area of the brain; rather, each recruits a broadly distributed neural network or processing stream. Moreover, certain areas of the brain that have been associated with language processing appear to be recruited across cognitive domains, suggesting that while language may be functionally special, it draws on at least some neural mechanisms and computational properties shared across other cognitive domains."

In order to claim that a particular brain region R supports a particular cognitive function, it is necessary not only to formulate predictions about the kinds of cognitive operations that should result in activity in region R, but also to be able to explain why other kinds of cognitive operations result in activity in region R. Researchers usually take one of two approaches (see e.g., a recent debate between Grodzinsky & Santi, 2008, and Willems & Hagoort, 2009, in TiCS, for an illustration): (1) they focus on the cognitive function that is of most interest to them (e.g., syntactic processing) and do not attempt to reconcile their findings with contradictory findings from other studies, sometimes dismissing them on the grounds of methodological flaws or slightly different loci of the peaks of activations; or (2) they attempt to provide a unifying account that can explain all the available data. For example, as mentioned above, some researchers have explicitly argued that the left inferior frontal brain regions that have been implicated in syntactic processing are more functionally general. For example, Hagoort and colleagues (e.g., Hagoort et al., 2009) have put forward

the *semantic unification* proposal, whereby the left IFG is responsible for constructing complex meanings out of simpler meaning units. Thompson-Schill and colleagues (Thompson-Schill, 2005; Novick et al., 2005; Snyder et al., 2007; see also Badre & Wagner, 2007) have argued that this region is critical for cognitive control abilities that are required to select the relevant representation among multiple competing representations. And Duncan (Duncan, 2001, 2010; see also Miller & Cohen, 2001) has argued that neurons located on the lateral surface of the frontal lobes are general-purpose problem solvers and can be used to perform a wide range of goal-directed cognitive tasks.

To summarize, no brain region has so far been convincingly demonstrated to be *selectively* engaged in syntactic processing or a component thereof. This may be because such regions do not exist. In particular, it is possible that all the regions that support syntactic processing support other, linguistic or non-linguistic, functions. However, we believe that current methods may be obscuring the true functional architecture, whatever that may be, as argued next.

Potential limitations of the current methods and a possible solution

All neuroimaging methods depend fundamentally on the ability to detect in some region of the brain a differential response between conditions that differ in the cognitive operations they involve. The current standard practice in the field of neuroimaging of language entails performing a random effects analysis on the data with the activations reported in stereotaxic coordinates (e.g., Talairach or MNI). The main way to compare results across studies is to compare the locations of activations (often represented by activation peaks) in stereotaxic space. The most systematic comparisons involve meta-analyses of activation peaks from large numbers of studies (e.g., Poldrack et al., 1999; Bookheimer et al., 2002; Indefrey & Levelt 2004; Kaan & Swaab, 2002; Costafreda et al., 2006; Vigneau et al., 2006; Lindenberg et al., 2007). The basic premise of these kinds of comparisons is that functional organization is sufficiently related to stereotaxic space, that this space provides a common ground for comparing results across studies, which is essential for accumulating knowledge about the functional architecture of the language system. However, because of inter-subject anatomical variability, functional regions may fall in different locations in stereotaxic space (see Saxe et al., 2006, and Fedorenko & Kanwisher, 2009, for reviews of the relevant literature and discussions of this and related issues). As a result, the responses across subjects extracted from the same location in stereotaxic space may underestimate selectivity of the underlying brain tissue.

Inter-subject anatomical variability has consequences for both voxel-based and standard-group-ROI-based analyses (where a fixed ROI, expressed as a set of voxels in stereotaxic space, is used across subjects; this includes both anatomically defined ROIs and functionally defined ROIs, or fROIs). In *voxel-based* analyses some subjects may not show an effect in a particular voxel. As a result, voxels that show an effect in only a small proportion of the subjects may not be detected, and in the voxels that do emerge, selectivity may be underestimated because averaging will include some subjects who do not show the effect (see also Appendix A). Similarly, in *standard-group-ROI-based* analyses individual subjects' activations may fall in slightly different locations relative to the ROI volume (which is fixed across subjects in common stereotaxic coordinates). As a result, selectivity may be underestimated because in each subject the ROI will include some voxels that do not show the effect.

The solution we offered in Fedorenko & Kanwisher (2009) involves functional localization of language-sensitive cortex in individual subjects (see Neville et al., 1998, Rogalsky & Hickok, 2000; Ben-Shachar et al., 2004; January et al., 2009, among others, for earlier attempts to use individual subjects' activation maps in the analyses). This approach consists

of two steps. First, a contrast aimed at the cognitive process of interest is used to identify regions – in each individual brain – that support the relevant cognitive process. For example, a contrast between faces and objects is used to identify regions that are engaged to a larger extent in face processing than in object processing more generally (Kanwisher et al., 1997). And second, hypotheses about the functions of these regions are tested by examining the response of these regions to various new conditions of interest.

In Fedorenko et al. (2010), we presented a localizer task that functionally identifies key language-sensitive regions in individual brains in a tractably short period of scanning (15–20 minutes). We used a contrast between sentences and pronounceable nonwords. This contrast targets brain regions engaged in processing word- and sentence-level meaning and should therefore include regions supporting lexico-semantic and syntactic processing. We demonstrated that a number of regions (including the "classical" regions in the left frontal and left temporo-parietal cortices) – each present in at least 80% of individual subjects – show a reliably stronger response to sentences than to nonwords. This response (a) replicates within and between subjects, (b) generalizes across different sets of materials and different tasks (see Osterhout et al., 2002, for evidence of task-independence in ERPs; cf. Meyer et al., 2000; Plante et al., 2002, Wright et al., 2010, among others, for evidence of task effects observed with linguistic materials in fMRI studies), and (c) generalizes from the visual to the auditory modality (see also Bedny et al., submitted), although some regions do show an overall stronger response to visual or auditory stimuli (see Fig. 8 in Fedorenko et al., 2010; see also Carpentier et al., 2001; Michael et al., 2001; Constable et al., 2004, for earlier investigations of modality effects in language processing; cf. Osterhout & Holcomb, 1993, for ERP evidence of similar responses to high-level linguistic manipulations across modalities). The ability to identify a set of regions in each individual brain in a short scan opens the door to a research program characterizing the response profiles of these regions in detail by examining their responses to various linguistic and non-linguistic manipulations, including investigations of questions concerning the role of these regions in syntactic processing (see Discussion for a discussion of advantages and potential limitations of this method).

In order to make a convincing case for this approach, however, it is necessary not only to demonstrate its feasibility as we have done in Fedorenko et al. (2010), but also to show how it compares to existing approaches. This is the aim of the current paper. In particular, we evaluate whether selectivity – which is critical for dissociating responses of a brain region to different stimuli/tasks – is indeed higher in subject-specific fROIs than in several kinds of standard group ROIs defined using traditional stereotaxic-space-based methods.

Selectivity

Throughout the paper, we use the term *selectivity* to characterize the degree to which different conditions are distinguished in a region's response. This represents a property of a *region*, and it is related to our ability to perform reverse inferences, i.e., inferring a particular cognitive process or computation from activation within a region (Poldrack, 2006). We also talk about the selectivity of a *method* when discussing the ability of a method to appropriately characterize brain regions in terms of stimuli/tasks that these regions respond to vs. do not respond to.

Formally, *selectivity* refers to the ability to detect an effect A while discriminating against a non-present effect B:

$$S = \beta_A \cdot (1 - \alpha_B) \quad \text{Equation 1}$$

where β_A represents the sensitivity (power) to detect effect A at a false positive level α. This selectivity measure represents the joint probability of detecting an effect A which is present (the true positive rate for effect A) while not detecting an effect B which is not present (the true negative rate for a null effect B). This selectivity measure is also known as the area under the curve (AUC), representing the area under the receiver operating characteristic (ROC) curve, and characterizing the *discrimination* ability of a statistical test (see Appendix A for a demonstration of how an increase in the size of the sample does not lead to an increase in selectivity, but rather decreases it).

As is clear from Equation 1 above, higher selectivity results from maximizing the likelihood of true positives (finding an effect when it is really present, i.e., *sensitivity* in the traditional statistics terminology) while minimizing the likelihood of false positives (not finding an effect when it is not present, i.e., *specificity* in the traditional statistics terminology).

We use effect sizes (i.e., percent BOLD signal differences) as an indicator of a method's selectivity. Effect sizes are important in understanding the functional architecture of a system, because they relate to the notion of "theoretical" significance (beyond the level of statistical significance). For example, if a region responds higher to one stimulus than another, but the difference between the means for these two conditions is very small, it is unclear how theoretically significant such a difference is (see Kanwisher, 2010, for a discussion of theoretical vs. statistical significance).

We chose to use raw (unstandardized) effect sizes for two reasons. First, standardized measures of effect size (such as Cohen's *d* or Hedge's *g*) are more commonly used when the metrics of the measures being studied do not have an intrinsic meaning. However, percent BOLD signal differences are meaningful in the sense that they are related to the amount of neural activity (despite many ambiguities about the precise kind of neural activity entailed). And second, even more importantly, differences in standardized measures have two possible interpretations: the differences could be due to (1) differences in (unstandardized) effect sizes (what we are reporting in the paper), and/or (2) differences in standard deviations of the data across the different methods. This latter difference is more difficult to interpret. In particular, we are arguing that the underlying BOLD signal differences for each subject are being underestimated when group (fixed) ROIs are used, due to inter-subject variability in the loci of activation, which leads to averaging – in any given subject – over voxels that show a strong effect and voxels that do not. Our direct comparison of the methods based on their unstandardized effect size contrast differences better reflects this underestimation thereby providing a stronger rationale for using subject-specific fROIs. (For completeness, we also provide the standardized effect size measures in Appendix G.)

Underestimating selectivity has important consequences for investigations of the functional architecture of the language system. If the method we are advocating has higher selectivity, then it has a potential to reveal functional architecture that may be invisible or diminished when less selective (i.e., group-based) methods are used. For example, underestimating the selectivity of regions supporting syntactic processing may lead to wrong conclusions about (a) within-language specificity (i.e., specificity for syntax vs. other aspects of language), (b) domain specificity (i.e., specificity for syntax vs. non-linguistic cognitive processes), and, consequently, (c) the linguistic (or other) computations conducted in these regions.

General methods

In this section, we (a) describe the dataset that we'll be using for all the method comparisons presented here, and (b) provide a summary of the individual-subjects functional localization

method that we recently developed. Other aspects of the methods, relevant to specific method comparisons, will be described in later sections.

The dataset

In the current paper we use a dataset (25 subjects) from a four-condition experiment, which used a blocked design and included sentences, word lists, jabberwocky sentences and pronounceable nonword lists (see e.g., Petersen et al., 1990; Mazoyer et al., 1993; Hagoort et al., 1999; Friederici et al., 2000b; Indefrey et al., 2001; Vandenberghe et al., 2002; Heim et al., 2005; Cutting et al., 2006; Humphries et al., 2006, 2007, for studies using similar contrasts in traditional group-based investigations)[2,3].

The sentences were 12 or 8 words long (two different sets of materials were used across two subsets of subjects to ensure generalizability across different sets of materials) and were created so as to include a variety of syntactic constructions. The word lists were created by scrambling the words across sets of sentences, so that "reconstructing" a sentence out of individual words in a word list was not possible. The jabberwocky sentences were created by replacing the content words (nouns, verbs, adjectives and adverbs) in the sentences by pronounceable nonwords that were matched to the words in length (in letters and syllables) and bigram letter frequency. Finally, the nonword lists were created by scrambling the nonwords across sets of jabberwocky sentences in a similar way to how words were scrambled in the word lists condition.

The largest contrast (sentences > nonwords) was used in Fedorenko et al. (2010) as the localizer contrast to define subject-specific functional ROIs. The word list and jabberwocky conditions were included in order to take a preliminary look at the question of whether any of the fROIs show a dissociation between syntactic and lexico-semantic processing.

The method: subject-specific functional ROIs

A detailed description of the method that we developed for defining subject-specific fROIs can be found in Fedorenko et al. (2010). We here provide a summary.

Because activations in many language tasks (including our localizer contrast) are quite distributed, the traditional subject-specific fROI definition procedure (e.g., Toottell et al., 1995; Kanwisher et al., 1997), where each individual map for the contrast of interest is examined and subsets of voxels are hand-selected, guided by macroanatomic landmarks, does not seem appropriate. Looking at the individual activation maps for various language tasks, it is often difficult to decide where the borders of a region are in any given brain, or what part of the activation counts as the "same" functional region across two brains. In order to reduce the potential subjectivity present in the traditional fROI definition procedure, we developed a novel method of fROI selection: Group-constrained Subject-Specific (GSS4) method.

In the GSS method, a set of thresholded activation maps from individual subjects are overlaid on top of one another. The topographical information in the resulting probabilistic

[2]The main results from this dataset are reported in Fedorenko et al. (2010). Detailed descriptions of the materials, procedure and other aspects of the experiment can be found there. The complete set of materials is available from http://web.mit.edu/evelina9/www/funcloc/.

[3]This dataset contains data from two experiments with the same design. In the first experiment (n=12) participants were engaged in a passive reading task, and in the second experiment (n=13), in addition to the reading task, participants had to report whether a memory probe (word/nonword) that appeared after each sentence/word list/jabberwocky sentence/nonword list was present in the immediately preceding string. Because the results from the two experiments were similar, we collapsed the data in Fedorenko et al. (2010) and in the current paper.

[4]This method was originally introduced with the abbreviation "GcSS".

overlap map is then used to divide the activation landscape into "partitions" (using an image parcellation algorithm) by identifying points of high inter-subject overlap and allowing for variation of individual activations around these high overlap points. In the critical step, the resulting group-level functional partitions are used to constrain the selection of voxels in individual subjects, i.e., individual subjects' fROIs are defined by taking the intersection of the partitions and individual subjects' thresholded activation maps. In the current version of the procedure, no constraints, such as e.g., contiguity, are placed on an individual subject's voxels that fall within the borders of any given partition (future work may include refinements to this procedure). Figure 1 illustrates this intersection procedure for a few sample subjects for two sample partitions (see http://web.mit.edu/evelina9/www/funcloc/funcloc_GcSS_analysisflow.html, for a schematic illustration of all the steps). Having intersected the partitions with individual subjects' activation maps, we selected a subset of "meaningful" fROIs that (1) are present in at least 80% of individual subjects, (2) show replicable responses within subjects and similar responses across subjects, and (3) respond more strongly to sentences than nonwords regardless of whether the stimuli are presented visually or auditorily (see Appendix B for a map showing the key group-level partitions identified in Fedorenko et al., 2010; these partitions, as well as an SPM toolbox for performing GSS analyses, are available for download from http://web.mit.edu/evelina9/www/funcloc/)5.

Unlike the traditional group analyses based on the notion of inter-subject overlap at the voxel level, this method enables us to detect individual subjects' activations that may land in similar but largely/completely non-overlapping locations within a spatially constrained volume.

Results: Selectivity of GSS fROIs vs. standard group ROIs

For all of these analyses, we use the dataset from Fedorenko et al. (2010), described in Methods. We extract the responses to the four conditions (sentences, words, jabberwocky and nonwords) from the different kinds of ROIs and compare the effect sizes observed in these ROIs for several key contrasts to those observed in our GSS fROIs (defined as described in Methods).

We focus on three contrasts relevant to syntactic processing: (1) sentences > nonwords (our localizer contrast), (2) sentences > words, and (3) jabberwocky > nonwords. The effect sizes for these contrasts tell us how well the relevant pairs of conditions are distinguished in a region's response (e.g., the size of the sentences > words effect tells us how well sentences and word lists are differentiated). Although none of these contrasts are designed to *isolate* syntactic processing exclusively, under any account of syntactic processing sentences should elicit a higher response than pronounceable nonwords and than words in regions supporting syntactic processing. Furthermore, according to most accounts, jabberwocky sentences should elicit a higher response than pronounceable nonwords because in the former but not in the latter, the order of nonwords, as well as the presence of function words in grammar-

[5]The brain regions identified with our localizer contrast are largely consistent with previous studies that have shown that left lateral frontal and temporal/temporo-parietal regions respond strongly to sentence stimuli, and also more strongly to more linguistically rich stimuli (e.g., sentences) than to less linguistically rich stimuli (e.g., word lists or jabberwocky) (e.g., Indefrey et al., 2001; Humphries et al., 2007). However, the results of one study (Friederici et al., 2000b) are not consistent with our findings. In particular, Friederici et al. (2000b) failed to observe stronger activation in the left frontal regions for sentences relative to word lists, jabberwocky sentences, and nonword lists. It is unclear what the source of this discrepancy is. Because we (see also Bedny et al., submitted) have now replicated the response profiles in the frontal and temporal regions across several studies (consistently finding the strongest response to sentences, weaker response to words and jabberwocky, and the weakest response to nonwords) and because these findings are also consistent with several other published reports, we think that the results reported in Friederici et al., 2000b are either not robust and therefore not replicated in similar studies, or are due to some study-specific properties which are different from all the other studies, although nothing stands out as an apparent difference based on a careful reading of the methods section of Friederici et al., 2000b.

appropriate locations and functional morphology allows combining nonwords into more complex representations6.

The goal of the current investigation is not to answer the question of whether any brain regions specialize for syntax vs. other linguistic processes and vs. non-linguistic functions. This is a hard question that will require dozens of studies. What we do here is establish the individual-subjects functional localization approach as a method that's optimally suited for answering this question eventually. In particular, we show this method's superior ability – relative to traditional, group-based, methods – to distinguish among conditions in a region's response (e.g., distinguishing between sentences and word lists). This ability is essential for addressing questions of functional specialization, which in turn lays the groundwork for understanding the computations performed by each brain region. (See Discussion for other advantages that the individual-subjects functional localization method provides over the traditional group-based methods.)

We begin, in Section A, by comparing the effect sizes in our subject-specific fROIs to those in standard group ROIs defined around activation peaks that have been argued to represent regions engaged in syntactic processing in several previously published studies. In Section B, we examine the effect sizes in our subject-specific fROIs in relation to fROIs defined based on the group analysis of the same data. Finally, in Section C, we examine two of our subject-specific fROIs (located in the left IFG) in relation to two anatomical ROIs: Brodmann areas (BAs) 44 and 45, defined using probabilistic cytoarchitectonic maps (Amunts et al., 1999). As discussed in the Introduction, the left IFG (including BAs 44 and 45) is among the regions most frequently implicated in syntactic processing and is therefore of relevance here.

A. GSS fROIs vs. ROIs defined around activation peaks from the literature

Methods—In order to perform a relatively large-scale assessment of the selectivity of syntax-sensitive regions reported in the literature, we chose to work with ROIs defined around activation peaks from a recent meta-analysis (Vigneau et al., 2006). Vigneau and colleagues examined activation peaks from a set of 129 studies. They classified these peaks (a total of 729 peaks) into three broad categories: peaks reported for tasks involving phonological processing, peaks reported for tasks involving semantic processing, and – of most interest here – peaks reported for tasks involving syntactic/sentence-level processing (see Vigneau et al., 2006, p. 1415, for a description of the classification procedure). 160 of the 729 peaks were classified as syntactic/sentence-level peaks. These 160 peaks were further classified into five sub-categories based on the type of contrast used: (1) text comprehension (42 peaks), (2) sentence-level comprehension (72 peaks), (3) syntactic processing (43 peaks), (4) sentence completion (1 peak), and (5) mental imagery (2 peaks). The most relevant here is the third subcategory, which, for the most part, includes contrasts between more vs. less structurally complex sentences. (The other four sub-categories are less theoretically coherent, including a wide range of tasks involving sentences and texts.)In particular, most of these studies investigated constructions that either did or did not contain long-distance dependencies between words (e.g., Stromswold et al., 1996; Cooke et al., 2002; Ben-Shachar et al., 2004; Constable et al., 2004). Long-distance dependencies have been shown to lead to processing difficulty, compared to local dependencies, across a wide

[6]It is worth noting that even if these conditions are met (i.e., the response to sentences is higher than to words and to nonwords, and the response to jabberwocky is higher than to nonwords), additional tests would then be required to better understand the role of the region in question in structural processing. Furthermore, in order to argue that a region is *selectively* engaged in syntactic processing, additional tests would be necessary to (a) demonstrate that this region is not engaged in other linguistic processes (e.g., phonological, semantic or discourse-level processes), and (b) demonstrate that this region is only sensitive to processing structure in language but not in other stimuli that may share some properties with language (e.g., music or action planning).

range of behavioral measures (e.g., Holmes & O'Regan, 1981; Ford, 1983; King & Just, 1991; Gibson, 1998; Gordon, Hendrick & Johnson, 2001). Because it is possible to match constructions involving local vs. non-local dependencies for lexico-semantic properties, contrasts between them are commonly used to isolate syntactic processing or syntactic working memory in behavioral, neuropsychological and neuroimaging investigations.

To perform the analyses in this section, we first (a) defined spherical ROIs (6mm in diameter) around each of the 43 syntactic peaks, and (b) extracted the response to the four conditions (sentences, words, jabberwocky and nonwords) from our dataset. We then selected the subset of these peaks that fell within one of our group-level partitions so that the data from these peaks would be comparable to the data from our new method. That constituted 29 peaks across 6 partitions (note that Vigneau et al. included only peaks in the left hemisphere in their meta-analysis). Finally, for each relevant partition, we compared the effect sizes in the ROIs defined around the syntactic peaks to those in the corresponding GSS fROI. We did this for each of our three contrasts: sentences > nonwords, sentences > words, and jabberwocky > nonwords. Before presenting the results, it is worth noting that these selectivity comparisons err on the conservative side because *all* the functional runs were used when estimating the responses in the peak-defined ROIs, whereas only the first run was used when estimating the response in GSS fROIs (the other runs were used for defining the ROIs).

Results—Across the three contrasts (sentences > nonwords, sentences > words, and jabberwocky > nonwords) and across different brain regions, the selectivity is higher in the fROIs defined with our GSS method than in nearby group ROIs defined around peak coordinates in stereotaxic space reported in earlier studies on syntactic processing. In particular, the differences in the magnitude of response to the conditions forming the contrasts above (e.g., the difference between the response to sentences and the response to word lists) are larger in the GSS fROIs compared to peak-defined group ROIs. The results are exemplified in Figure 2 for two sample regions (see Appendix C, for a full set of detailed plots).

To quantify these differences in effect sizes, we used a Wilcoxon signed rank one-sided test comparing the GSS fROI effect size value to the corresponding distribution of peak-defined ROIs' effect size values. The results are summarized in Table 1. In three of the regions (IFG, IFGorb and MidPostTemp) the GSS fROI effect size is reliably higher than those of the peak-defined ROIs across all three contrasts. In the PostTemp region, the GSS fROI effect size is marginally higher than those of the peak-defined ROIs across the three contrasts. And in the remaining two regions (MFG and AntTemp) the effect size is only numerically higher, plausibly not reaching significance due a small number of peak-defined ROIs in those partitions.

In summary, these results demonstrate that GSS fROIs are more functionally selective than ROIs defined around activation peaks reported for syntactic contrasts in the previous literature. This was true across all three contrasts that involve syntactic processing. These results therefore suggest that analyses relying on GSS fROIs are promising for investigating questions of the functional architecture of the language system as they have more power to detect effects if they are present and to distinguish among conditions that are actually different.

B. GSS fROIs vs. standard group fROIs (defined in the same dataset)

Methods—In this section, we compare the selectivity of our GSS fROIs to that of fROIs defined based on traditional random-effects group analysis maps in the same subjects. For both kinds of ROIs, we used all but the first functional run in each subject to define the

ROIs, and then extracted the responses from the first functional run, i.e., from a subset of the data that was not used for ROI definition (e.g., Kriegeskorte et al., 2009; Vul & Kanwisher, 2009). To define our GSS fROIs, we intersected group-level partitions (see Appendix B) with individual subjects' activation maps for the sentences > nonwords contrast, as discussed in Methods and illustrated – for two sample partitions and six sample subjects – in Figure 1 above. To define standard group fROIs, we intersected group-level partitions with the random-effects group analysis map for the same contrast (again, using all but the first functional run). Critically, in the group fROI analyses the same set of voxels in stereotaxic space was used for all subjects (i.e., the voxels that emerged in the group analysis for the sentences > nonwords contrast and fell within the borders of each relevant partition), whereas our method allows different voxels in stereotaxic space for each individual subject.

As discussed in Fedorenko et al. (2010), it is not obvious that the GSS fROIs would outperform group fROIs in this comparison. In particular, because the random-effects group fROIs include voxels that behave most consistently across participants, it is possible that these voxels would be most functionally selective. Therefore, the selectivity of group fROIs may be predicted a priori to be at least as high as that of subject-specific fROIs. This is not, however, what the results show. In Fedorenko et al. (2010) we already reported the results for the sentences > nonwords contrast. For each of the twelve regions we observed a reliably higher selectivity for the GSS fROIs compared to the group fROI. Here, we examined the selectivity of the GSS fROIs and group fROIs for the two functionally narrower contrasts that are relevant to syntactic processing: sentences > words, and jabberwocky > nonwords.

Results—For each of the three contrasts across brain regions, the selectivity was higher in GSS fROIs than in group fROIs (p-values <.0001). Looking at the individual regions, we see that numerically the effect is larger in GSS fROIs, compared to group fROIs, in every single region across the three contrasts. These results are presented in Figure 3 (the results for the sentences > nonwords contrast are the same as those reported in Fedorenko et al., 2010, and are presented here for completeness). Moreover, in most of the regions, these differences in effect size values between GSS fROIs and group fROIs are significant (for example, we see a highly reliable difference across all three contrasts in the Left IFG and Left MidPostTemp regions). These results are summarized in Table 2. These results reflect the fact that in group ROIs, for each subject the averaging includes some voxels that are not significant in that subject, due to inter-subject variability in the precise locations of activations.

Although the contrasts discussed here (sentences > nonwords, sentences > words, and jabberwocky > nonwords) are *significant* in most of the group ROIs, it is important to understand that smaller effect sizes mean that it would be generally more difficult to detect these (and other similar-size) effects. For example, while our dataset includes 25 subjects, typical imaging experiments where standard group analyses are used include on the order of 14–18 subjects. Therefore, some of these effects may not be detectable in such studies.

It is also worth noting that apart from the fact that GSS fROIs are characterized by higher selectivity, there is some variability among the regions in how selective they are (this variability is present in both GSS fROIs and group fROIs, but it is more pronounced in GSS fROIs, plausibly because the selectivity is overall higher allowing for more variation). Given how important selectivity of different brain regions is for understanding the neural architecture underlying human cognition, an interesting future line of research could be aimed at exploring this variability. In particular, understanding what properties of a region (such as e.g., its size, location in the brain, or spatial variability across subjects) predict the size of the effects observed in that region for various tasks may provide important insights. (For example, it is possible that language-sensitive regions located in the left hemisphere

would show consistently larger effect sizes for language tasks than those in the right hemisphere.)In Appendix D we present some preliminary explorations of this question.

C. GSS fROIs vs. anatomical ROIs (BAs 44 and 45)

In this section, we compare two of our left frontal GSS fROIs (Left IFG and Left IFGorb) to two anatomical ROIs. In particular, we investigate Brodmann Areas 44 and 45, defined using probabilistic cytoarchitectonic maps (Amunts et al., 1999). These two Brodmann areas constitute Broca's region, which is one of the main candidates for the seat of syntactic processing, as discussed in the Introduction.

Methods—We first discuss the methodology of using probabilistic cytoarchitectonic maps more generally, since this tool is relatively new to the field of neuroimaging of language. Then, we discuss the relationship between BAs 44 and 45 and our two GSS fROIs located in the left IFG.

Probabilistic cytoarchitectonic maps: Although locations of cytoarchitectonic areas (typically Brodmann areas) are commonly estimated in neuroimaging studies based on the sulcal and gyral anatomy, these estimates are imperfect because of high inter-subject variability in the locations of different cytoarchitectonic zones relative to macroanatomic landmarks (e.g., Zilles et al., 1997; Amunts et al., 1999; Amunts & Willmes, 2006). Several years ago, Zilles and colleagues (Schormann and Zilles, 1998; Amunts et al. 1999; Zilles et al. 2002) embarkedon an ambitious project where the borders of cytoarchitectonic areas were delineated in ten post-mortem human brains. These brains were first scanned at high resolution. Then, each brain was thinly sliced and each target cytoarchitectonic zone was delineated in the relevant set of slices in each brain. Finally, the delineated sections in the relevant slices were aligned with the MR volume of the brain. Normalizing these ten brains into a common space then enabled the researchers to create probabilistic cytoarchitectonic maps, i.e. representations of the percent of subjects who show a given Brodmann area in each voxel of the normalized (MNI) space. So far, thirteen Brodmann areas have been delineated, and probabilistic maps for these areas are now publicly available. Among the first were Brodmann areas 44 and 45 (Amunts et al., 1999).

These cytoarchitectonic maps are a powerful tool as they bring us one step closer to being able to detect structure below the level of sulci/gyri in fMRI. Some researchers (e.g., Grodzinsky, 2010; cf. Fedorenko & Kanwisher, in press) have argued that these maps are an ideal way to answer questions of functional specialization in the human brain, because (a) they represent cytoarchitecture which has been long assumed (e.g., Brodmann, 1909) and, based on the animal work, shown to correspond to function (e.g., Iwamura et al., 1983; Matelli et al., 1991; Rozzi et al., 2008), and (b) they contain information about the variability present across individual brains.

We agree that these maps provide an excellent tool for use in fMRI, allowing a much more precise estimation of the location of cytoarchitecture relative to macroanatomy. However, we also think that some of the problems inherent in the traditional group analyses affect these maps, as well. In particular, due to large variability in the locations of these areas across subjects, maps for adjacent areas exhibit substantial overlap (for example, 43.0% of BA44 voxels also belong to BA45, and 51.3% of BA45 voxels also belong to BA44). As a result, deciding which cytoarchitectonic area a given voxel belongs to is non-trivial. It can be done using high thresholding levels for the cytoarchitectonic maps, i.e., including – for each BA – only voxels that are present in most of the 10 brains7. However, this thresholding inevitably leads to a loss of information about the variability in the underlying anatomy, which is what these maps are trying to capture in the first place.

An alternative is to calculate for a target functional cluster of interest the probability that any given voxel in that cluster would be labeled as belonging to a particular cytoarchitectonic area (e.g., Santi & Grodzinsky, 2007). This is done by (a) extracting – for each voxel in the functional cluster – the probability of that voxel being in a given Brodmann area, and (b) averaging the values across all the voxels in the cluster. This approach provides useful information about the degree to which functional activations land within a particular Brodmann area. However, no standards have yet been set for what probability values are sufficiently high in order to claim that a functional cluster lands within the borders of a particular Brodmann area. Setting an absolute stringent threshold seems too arbitrary, so quantitative methods for this kind of an assessment need to be developed. These methods would presumably take into account the size of the functional cluster, the variability in the locations of individual Brodmann areas for the target area, the extent of overlap between the target Brodmann area and adjacent Brodmann areas, the topography of the functional cluster relative to the topography of the target Brodmann area, etc. For example, given the substantial overlap between areas BA44 and BA45, it may be necessary to determine the minimum difference in probability values between these two areas that would be convincing in claiming that a functional cluster corresponds to e.g., BA45 and not BA44. The degrees of freedom currently present in this process make any claims about the relationship between functional activations and these probabilistic cytoarchitectonic areas difficult to interpret. (see Eickhoff, Paus, Caspers, Grosbras, Evans, Zilles, and Amunts, 2007, for a discussion of issues in assigning activations to cytoarchitectonic areas and a possible solution)

BA44 and BA45 vs. Left IFG and Left IFGorb fROIs: We here focus on two of our GSS fROIs: the ones that overlap with the two relevant Brodmann areas. Figure 4 demonstrates the relationship between the group-level partitions corresponding to our IFG regions (Left IFG and Left IFGorb) and BA44 and BA45. As can be seen from the figure, our left IFG partition overlaps with both BA44 and BA45 (which overlap with each other considerably, as discussed above), and the left IFGorb partition overlaps with BA45, but almost not at all with BA44 (see Appendix E for more information on the degrees of overlap).

In comparing the selectivity of the Brodmann areas to that of our GSS fROIs, we used (a) all the voxels in each BA map, and (b) only the voxels that are present in nine out of ten[8] brains in Amunts et al.'s analysis (i.e., only high-overlap voxels)[9] in order to see whether the results might look clearer at a higher threshold. In addition, we intersected the probabilistic maps with individual subjects' activation maps (again, using all but the first functional run) and defined subject-specific ROIs in the same way as we did using group-level partitions in our GSS method. This was done in an attempt to examine selectivity in a different kind of *subject-specific* fROIs. As discussed in Fedorenko et al. (2010), group-level functional partitions are only one way to constrain the selection of subject-specific voxels. Other possibilities are to constrain subject-specific ROIs with boundaries of macroanatomical structures with respect to each individual subject's anatomy (the current standard in the fROI method in high-level vision studies), boundaries of macroanatomical structures derived from probabilistic atlases (e.g., Westbury et al., 1999;Rademacher et al., 2001), or functional constraints derived in some different way (e.g., by using the borders of a fROI reported in

[7]For reference, here is information on the overlap between BA 44 and BA45 across different thresholds (the first number represents the percentage of BA44 voxels that also belong to BA45, and the second number represents the percentage of BA45 voxels that also belong to BA44; n is the threshold in terms of the minimal number of subjects for whom a voxel corresponds to the relevant BA): n>0: 43.0, 51.3; n>1: 41.4, 46.4; n>2: 27.6, 40.7; n>3: 25.3, 25.4; n>4: 20.5, 17.8; n>5: 12.6, 6.9; n>6: 4.1, 1.7; n>7: 0.4, 0.0; n>8: 0.0, 0.0; n>9: NA, 0.0.

[8]Because in BA44 there is no single voxel where all ten subjects overlap (in the MNI space; see Fischl et al., 2008, for better alignment results in the spherical space), we used the 9/10 threshold for both BAs.

[9]Because the responses to the four conditions were similar across different thresholds (see Appendix F), we only considered the two extreme thresholds here.

another study). Using probabilistic cytoarchitectonic maps is yet another possibility (modulo the limitation that such maps only exist for a subset of cytoarchitectonic areas so far).

In order to make the results comparable across these different kinds of ROIs, even for the extraction of responses from the BA ROIs (where no data were used for defining the ROIs), we used the data from the first run only.

Results—The results for the three contrasts (sentences > nonwords, sentences > words, and jabberwocky > nonwords) are presented in Figure 5.

Across all three contrasts the effect sizes are much larger in either of the two GSS fROIs (5b), compared to either way of defining the anatomical ROIs based on the probabilistic cytoarchitectonic maps (all voxels or just high-overlap voxels), 5a. In the comparison between the *Left IFG* GSS fROI and (i) BA 44 all voxels, (ii) BA 45 all voxels, (iii) BA 44 high-overlap voxels, and (iv) BA 45 high-overlap voxels, we see reliably larger effects for the sentences > nonwords contrast (p-values: <.0001, .0001, .001, .0001), for the sentences > words contrast (p-values: <.0001, .0001, .05, .001), and for the jabberwocky > nonwords contrast (p-values: <.0001, .001, .01, .001). Similarly, in the comparison between the *Left IFGorb* GSS fROI and (i) BA 44 all voxels, (ii) BA 45 all voxels, (iii) BA 44 high-overlap voxels, and (iv) BA 45 high-overlap voxels, we see reliably larger effects for the sentences > nonwords contrast (p-values: <.0001, .0001, .01, .0001), for the sentences > words contrast (p-values: <.001, .001, .01, .01), and for the jabberwocky > nonwords contrast (p-values: <. 001, .001, .01, .001).

In addition, setting the differences in selectivity aside, it is worth noting that although some of the syntactic contrasts are significant in the BA ROIs (e.g., the difference between sentences and nonwords is significant in each of the BAs in the analysis with all the voxels (in both areas p<.02) and in BA45 in the analysis of high-overlap voxels (p<.02)), these significance levels would not survive any reasonable correction for multiple comparisons. (See Appendix F for response magnitudes for the four conditions and a further discussion of the response profiles of the BA ROIs.)

Is the higher power of the GSS fROIs compared to the BA ROIs due to a difference between our group-level functional partitions and the BA maps, or is it due to the importance of intersecting any kind of spatial constraint (functional or anatomical) with individual subjects' activation maps? In Figure 5c we show the results for ROIs defined by intersecting individual subjects' activation maps with the BA maps. The selectivity or these ROIs is comparable to that observed for GSS fROIs (in fact, none of the differences in effect sizes between the GSS fROIs and the subject-specific ROIs defined using the BA maps are significant). This result suggests that it is not the specific partitions that are important for selectivity but rather the intersection of these partitions with each individual subject's localizer activations.

Summary of Sections A-C

In Sections A-C, we compared the selectivity of our GSS fROIs to different kinds of standard group ROIs where the same set of voxels in stereotaxic coordinates is used across subjects, and we have shown that GSS fROIs indeed have higher selectivity than the ROIs in each of the alternative methods considered. This strengthens our claim that the individual-subjects functional localization approach is optimally suited for investigating questions of functional specialization and holds promise for uncovering functional architecture in the language system that may have been invisible to other methods. And to the extent that we discover – in future work – overlap between e.g., syntactic processing and other aspects of language, or between syntactic processing and some non-linguistic task, using this more

selective method, we will be more confident that what we have discovered is true multi-functionality, rather than a false positive resulting from averaging across adjacent but functionally distinct regions.

Discussion

The goal of this paper was to test whether the individual-subjects functional localization method not only presents a viable alternative to the traditional methods used in the neuroimaging of language community, but whether it is superior to those methods in its ability to differentiate among conditions in language-sensitive brain regions. We directly compared the individual-subjects functional localization approach and several traditional approaches, focusing on regions implicated in syntactic processing. The GSS fROIs exhibited higher selectivity than (a) ROIs defined around the activation peaks from previous studies that have claimed to identify syntax-sensitive regions, (b) fROIs defined on the group analysis maps of the same data, and (c) anatomical ROIs defined using probabilistic cytoarchitectonic maps.

As discussed in the Introduction, important questions about the neural architecture supporting syntactic processing currently remain unanswered. These questions concern the nature and degree of functional specialization of brain regions for syntactic processing vs. (a) other aspects of linguistic processing, and (b) non-linguistic cognitive tasks that may involve similar representations or computations. Answers to these questions are of interest in their own right, and further provide a necessary foundation for the ultimate goal of characterizing the computations conducted in these regions. Although we have not yet answered these questions, we have here demonstrated that the standard stereotaxic-space-based methods may not be ideally suited for investigating questions of functional specificity. In particular, in order to uncover potential functional dissociations, the most sensitive available method should be used, such that the chances of revealing such dissociations would be maximized, while the chances of observing spurious overlap would be minimized. We have shown that the traditional, group-based, methods underestimate selectivity, compared to the individual-subjects functional localization approach. We therefore conclude that functional localization in individual subjects holds promise for future investigations of syntactic processing, as well as for the domain of language more generally. In particular, by investigating the responses of language-sensitive regions – defined in individual subjects – to linguistic and non-linguistic manipulations, we can arrive at a detailed characterization of each region and start generating hypotheses about the kinds of representations each region stores or manipulates and the kinds of computations it performs.

In addition to demonstrating higher selectivity of our GSS fROIs, relative to several kinds of standard group ROIs, we wanted to elaborate more generally on the informativeness/usefulness of activation peaks reported in stereotaxic space given that we are making an argument for the need to define regions under investigation functionally in individual subjects. Perhaps the most obvious way to assess the usefulness of activation peaks in stereotaxic space would be to examine their *replicability*. This could be done by selecting a small set of studies from the literature and attempting to replicate their findings using the same exact experimental protocol and the same scanning and analysis procedures, in order to examine the locations of activations across different versions of the same study. Activation peaks in stereotaxic space only have value to the extent that they represent some stable property of the surrounding neural tissue in that space. If this is the case, then an exact replication of a study should result in activations landing in very similar stereotaxic locations. If differences in the precise locations of activations across studies are sufficiently large, this already poses a problem for the current methods. In particular, given that the extent of this kind of variability is unknown in most cases10 (or at least is not quantified), it

is unclear whether activations reported, for example, for two different tasks and landing in similar locations should be interpreted as activity of the same functional region or distinct adjacent regions. Without formally established criteria for such decisions, researchers are in danger of being affected by their theoretical biases in interpreting such patterns of activation.

An alternative way to evaluate the informativeness of activation peaks reported in stereotaxic space is to examine the responses of ROIs defined around these peaks to some conditions in new datasets. Such analyses would assess the *generalizability* of the peaks. The logic is as follows: if a particular location in stereotaxic space (e.g., an activation peak) represents a region that supports some cognitive function, then neural tissue surrounding this location should respond to various new stimuli/tasks in ways consistent with the hypothesized function. For example, it is reasonable to expect that a region argued to support auditory processing (based on e.g., a contrast between listening to tones and looking at gratings) would respond more strongly to music or speech sounds compared to e.g., pictures of objects. Because we already had the data on the response of the ROIs defined around the peaks meta-analyzed by Vigneau et al. to the four conditions in our dataset (sentences, words, jabberwocky and nonwords), we took a preliminary look at this question. The results were disappointing. For example, among the syntactic (n=43) and sentence-level (n=72) ROIs only a little over half respond to sentences reliably stronger than to nonwords/words at the most liberal threshold. This result is disturbing especially given that the contrast between sentences and nonwords/words is quite large (cf. a contrast between more and less complex sentences). If neural tissue around these peaks does not show even the most rudimentary characteristics that would be expected of regions that these peaks are argued to represent, then how can we treat these peaks as meaningful and rely on them in comparing results across studies, which is essential for accumulating knowledge about the functional profiles of language-sensitive cortex? In order to build on the results of previous studies, we need some reason to believe that the same functional regions are being characterized across studies. The GSS fROI approach provides one way to do this.

Note that in advocating the use of fROIs defined in each subject individually, for the study of the functional architecture of the language system, we are *not* arguing that other methods, or data based on them, should be discarded. Rather we see the individual fROI method as complementary in many ways to the standard group analyses (see Fedorenko et al., 2010, for additional discussion). For one thing, the construction of the best localizers (still very much work in progress) relies on some form of group analyses to find the regions characterized by highest inter-subject overlap. In addition, functional dissociations found with group analyses – as long as they are replicable – remain very powerful, as these phenomena are found despite the tendency to underestimate selectivity in group analyses.

Before concluding, it is important to point out some limitations inherent in the individual-subjects functional localization approach and to discuss how these limitations can be addressed (see also Fedorenko et al., 2010). There are two primary concerns with interpreting the results obtained using the fROI method: (1) coverage, and (2) response profile ambiguity. First, it is possible that the localizer selected to identify brain regions of interest is not inclusive enough and some key region(s) are left out. With respect to this issue, our localizer includes many of the regions previously implicated in linguistic processing. As a result, even though it is possible that we are not including some important language-sensitive regions, a detailed characterization of the ones that we do capture will

[10]Because of the expense associated with fMRI, there are very few “pure” replications, which would be important for quantifying the variability in the precise locations of activations for the same functional contrasts. (Critically, replication attempts would help eliminate spurious findings from consideration and strengthen our beliefs in a subset of findings that have a real empirical basis.)

significantly advance our understanding of how language is implemented in the brain. (Besides, whole-brain analyses, which should always be used to supplement fROI-based analyses, will tell us if some systematic activations outside the borders of the fROIs are missed. If such activations are observed, additional new fROIs should be considered.)

And second, response profiles of functionally defined ROIs are often ambiguous. For example, a response profile where an ROI responds similarly strongly to two conditions is consistent with both (a) true multi-functionality, and (b) functionally distinct subregions within the ROI that respond to each of the conditions. Similarly, a response profile where an ROI does not respond to some condition could be observed when only a small proportion of voxels in that ROI respond to that condition. These concerns can be addressed by supplementing the fROI analyses with whole-brain analyses. Whole-brain analyses may detect activations outside the borders of the fROIs as well as non-homogeneous activations within the fROIs. Furthermore, in our recent work we developed analysis tools that enable examining subsets of voxels within individual subjects' fROIs that may have distinct functional profiles (these analysis tools are currently available by request and will soon be made available from http://web.mit.edu/evelina9/www/funcloc.html). In cases where some functionally distinct subsets of voxels are discovered within a fROI, new fROI definitions should be considered that would divide some of the fROIs into smaller regions of analysis.

Coming back to the question of the neural basis of syntactic processing, although we remain agnostic as to whether any machinery exists in the human brain that is specialized for processing structural aspects of the linguistic signal, we believe that traditional methods – used in the vast majority of previous studies – have faced important limitations in providing a clear answer to this and related questions. In particular, group-based investigations (a) underestimate selectivity by not taking into consideration variability in activation loci across subjects; (b) are rarely replicated making it difficult to decide which findings are real and which may be spurious; and (c) do not provide an easy way to compare results across studies, thereby slowing down the process of knowledge accumulation about the functional profiles of key language-sensitive regions. The individual-subjects functional localization approach circumvents these problems. In addition to its superior ability to distinguish among conditions in a brain region's response, as demonstrated in the current paper (see also Saxe et al., 2006), the subject-specific fROI approach (i) always includes a partial replication of the results of previous studies (by showing similar responses to the localizer conditions in a subset of the data not used for ROI definition), and (ii) provides a reliable way to refer to the same functional region across individuals and studies, leading to faster progress in discerning the function(s) of brain regions.

To conclude, maybe all language-sensitive regions indeed support multiple aspects of language and/or some non-linguistic cognitive operations (e.g., Blumstein, 2009). If so, we will find stronger evidence of such multi-functionality and will be able to better characterize it by examining overlap effects at the individual-subject level. It is also possible, however, that stunning functional specificity has simply remained hidden in previous studies. If so, the approach advocated here stands a good chance of finding it.

Acknowledgments

We would like to thank members of the Kanwisher and Gibson labs for comments on the analyses reported here, and Sarah Weigelt and five anonymous reviewers for comments on the earlier drafts of the manuscript. We are also grateful to Mathieu Vigneau and colleagues for providing us with the entire set of activation peaks from Vigneau et al. (2006) with the relevant descriptions. This research was supported by a grant from the Ellison Medical Foundation to NK and by award number K99HD057522 from the Eunice Kennedy Shriver National Institute Of Child Health & Human Development to EF. We also would like to acknowledge Athinoula A. Martinos Imaging Center at McGovern Institute for Brain Research, MIT.

References

Amunts K, Willmes K. From inter-subject variability in clinical syndromes to anatomical variability. Brain Lang. 2006; 96(2):147–50. discussion 157–70. [PubMed: 16085299]

Amunts K, Schleicher A, Bürgel U, Mohlberg H, Uylings HB, Zilles K. Broca's region revisited: Cytoarchitecture and intersubject variability. J Comp Neurol. 1999; 412(2):319–41. [PubMed: 10441759]

Badre D, D'Esposito M. Is the rostro-caudal axis of the frontal lobe hierarchical? Nature Reviews Neuroscience. 2009; 10:659–669.

Badre D, Wagner AD. Left ventrolateral prefrontal cortex and the cognitive control of memory. Neuropsychologia. 2007; 45(13):2883–2901. [PubMed: 17675110]

Bedny, M.; Dodell-Feder, D.; Pascual-Leone, A.; Fedorenko, E.; Saxe, R. Language processing in the occipital cortex of congenitally blind adults; Proceedings of the National Academy of Sciences of the United States of America; in press

Ben-Shachar M, Hendler T, Kahn I, Ben-Bashat D, Grodzinsky Y. The neural reality of syntactic transformations: Evidence from functional magnetic resonance imaging. Psychol Sci. 2003; 14(5): 433–40. [PubMed: 12930473]

Ben-Shachar M, Palti D, Grodzinsky Y. Neural correlates of syntactic movement: Converging evidence from two fmri experiments. Neuroimage. 2004; 21(4):1320–36. [PubMed: 15050558]

Binkofski F, Amunts K, Stephan KM, Posse S, Schormann T, Freund HJ, et al. Broca's region subserves imagery of motion: A combined cytoarchitectonic and fmri study. Hum Brain Mapp. 2000; 11(4):273–85. [PubMed: 11144756]

Blumstein, SE. The cognitive neurosciences. IV. Cambridge, MA: MIT Press; 2009. Reflections on the cognitive neuroscience of language.

Blumstein SE, Myers EB, Rissman J. The perception of voice onset time: An fmri investigation of phonetic category structure. J Cogn Neurosci. 2005; 17(9):1353–66. [PubMed: 16197689]

Bokde AL, Tagamets MA, Friedman RB, Horwitz B. Functional interactions of the inferior frontal cortex during the processing of words and word-like stimuli. Neuron. 2001; 30(2):609–17. [PubMed: 11395018]

Bornkessel-Schlesewsky, I.; Friederici, A. Neuroimaging studies of sentence and discourse comprehension. In: Gaskell, editor. The oxford handbook of psycholinguistics. United Kingdom: Oxford Univ Press; 2007. p. 407-24.

Bookheimer SY. Functional MRI of Language: new approaches to understanding the cortical organization of semantic processing. Annual Review of Neurosciences. 2002; 25:151–88.

Brodmann, K. Vergleichende lokalisationslehre der grosshirnrinde in ihren prinzipien dargestellt auf grund des zellenbaues. Barth Leipzig; 1909.

Buckner RL, Raichle ME, Petersen SE. Dissociation of human prefrontal cortical areas across different speech production tasks and gender groups. J Neurophysiol. 1995; 74(5):2163–73. [PubMed: 8592204]

Burton MW. The role of inferior frontal cortex in phonological processing. Cognitive Science. 2001; 25(5):695–709.

Burton MW. Understanding the role of the prefrontal cortex in phonological processing. Clin Linguist Phon. 2009; 23(3):180–95. [PubMed: 19283576]

Caplan D. Functional neuroimaging studies of syntactic processing in sentence comprehension: A critical selective review. Language and Linguistics Compass. 2007; 1(1–2):32–47.

Carramazza A, Zurif E. Dissociations of algorithmic and heuristic processes in sentence comprehension: Evidence from aphasia. Brain Lang. 1976; 3:572–582. [PubMed: 974731]

Carpentier A, Pugh KR, Westerveld M, Studholme C, Skrinjiar O, Thompson JL, Spencer DD, Constable RT. Functional MRI of Language Processing: Dependence on Input Modality and Temporal Lobe Epilepsy. Epilepsia. 2001; 42(10):1241–1254. [PubMed: 11737158]

Constable RT, Pugh KR, Berroya E, Mencl WE, Westerveld M, Ni W, et al. Sentence complexity and input modality effects in sentence comprehension: An fmri study. Neuroimage. 2004; 22(1):11–21. [PubMed: 15109993]

Cooke A, Zurif EB, DeVita C, Alsop D, Koenig P, Detre J, et al. Neural basis for sentence comprehension: Grammatical and short-term memory components. Hum Brain Mapp. 2002; 15(2): 80–94. [PubMed: 11835600]

Costafreda SG, Fu CH, Lee L, Everitt B, Brammer MJ, David AS. A systematic review and quantitative appraisal of fmri studies of verbal fluency: Role of the left inferior frontal gyrus. Hum Brain Mapp. 2006; 27(10):799–810. [PubMed: 16511886]

Cutting LE, Clements AM, Courtney S, Rimrodt SL, Schafer JG, Bisesi J, et al. Differential components of sentence comprehension: Beyond single word reading and memory. Neuroimage. 2006; 29(2):429–38. [PubMed: 16253527]

Dapretto M, Bookheimer SY. Form and content: Dissociating syntax and semantics in sentence comprehension. Neuron. 1999; 24:427–432. [PubMed: 10571235]

Devlin JT, Matthews PM, Rushworth MF. Semantic processing in the left inferior prefrontal cortex: A combined functional magnetic resonance imaging and transcranial magnetic stimulation study. J Cogn Neurosci. 2003; 15(1):71–84. [PubMed: 12590844]

Duncan J. An adaptive coding model of neural function in prefrontal cortex. Nature Reviews Neuroscience. 2001; 2(11):820–829.

Duncan J. The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in Cognitive Sciences. 2010; 14(4):172–179. [PubMed: 20171926]

Eickoff SB, Paus T, Caspers S, Grosbras MH, Evans A, Zilles K, Amunts K. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. Neuroimage. 2007; 36:511–521. [PubMed: 17499520]

Fadiga L, Craighero L, D'Ausilio A. Broca's area in language, action, and music. Ann N Y Acad Sci. 2009; 1169:448–58. [PubMed: 19673823]

Fedorenko E, Kanwisher N. Neuroimaging of language: Why hasn't a clearer picture emerged? Language and Linguistics Compass. 2009; 3(4):839–865.

Fedorenko E, Kanwisher N. Functionally localizing language-sensitive regions in individual subjects with fMRI: A reply to Grodzinsky's critique of Fedorenko & Kanwisher 2009. Language and Linguistics Compass. 2011; 5(2):78–94.

Fedorenko E, Hsieh PJ, Nieto-Castanon A, Whitfield-Gabrieli S, Kanwisher N. A new method for fMRI investigations of language: Defining ROIs functionally in individual subjects. Journal of Neurophysiology. 2010; 104:1177–1194. [PubMed: 20410363]

Fiez JA. Phonology, semantics, and the role of the left inferior prefrontal cortex. Hum Brain Mapp. 1997; 5(2):79–83. [PubMed: 10096412]

Fischl B, Rajendran N, Busa E, Augustinack J, Hilds O, Thomas Yeo BT, Mohlberg H, Amunts K, Zilles K. Cortical folding patterns and predicting cytoarchitecture. Cerebral Cortex. 2008; 18(8): 1973–1980. [PubMed: 18079129]

Ford M. A method for obtaining measures of local parsing complexity throughout sentences. Journal of Verbal Learning and Verbal Behavior. 1983; 22:203–218.

Friederici AD, Wang Y, Herrmann CS, Maess B, Oertel U. Localization of early syntactic processes in frontal and temporal cortical areas: A magnetoencephalographic study. Human Brain Mapping. 2000a; 11:1–11. [PubMed: 10997849]

Friederici AD, Meyer M, von Cramon DY. Auditory language comprehension: An event-related fmri study on the processing of syntactic and lexical information. Brain Lang. 2000b; 75(3):289–300. [PubMed: 11386224]

Friederici AD, Kotz SA. The brain basis of syntactic processes: Functional imaging and lesion studies. Neuroimage. 2003; 20:8–17.

Friederici AD, Ruschenmeyer SA, Hahne A, Fiebach CJ. The role of left inferior frontal and superior temporal cortex in sentence comprehension: Localizing syntactic and semantic processes. Cerebral Cortex. 2003; 13:170–177. [PubMed: 12507948]

Friederici AD, Fiebach CJ, Schlesewsky M, Bornkessel ID, von Cramon DY. Processing linguistic complexity and grammaticality in the left frontal cortex. Cereb Cortex. 2006; 16(12):1709–17. [PubMed: 16400163]

Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC. Recursive syntactic pattern learning by songbirds. Nature. 2006; 440(7088):1204–7. [PubMed: 16641998]

Gibson E. Linguistic complexity: Locality of syntactic dependencies. Cognition. 1998; 68:1–76. [PubMed: 9775516]

Gordon P, Hendrick R, Johnson M. Memory interference during language processing. Journal of Experimental Psychology: Learning, Memory, & Cognition. 2001; 27(6):1411–1423.

Gough PM, Nobre AC, Devlin JT. Dissociating linguistic processes in the left inferior frontal cortex with transcranial magnetic stimulation. J Neurosci. 2005; 25(35):8010–6. [PubMed: 16135758]

Graves WW, Grabowski TJ, Mehta S, Gupta P. The left posterior superior temporal gyrus participates specifically in accessing lexical phonology. J Cogn Neurosci. 2008; 20(9):1698–710. [PubMed: 18345989]

Grewe T, Bornkessel I, Zysset S, Wiese R, von Cramon DY, Schlesewsky M. The emergence of the unmarked: A new perspective on the language-specific function of Broca's area. Hum Brain Mapp. 2005; 26(3):178–90. [PubMed: 15929098]

Grodzinsky Y, Friederici AD. Neuroimaging of syntax and syntactic processing. Curr Opin Neurobiol. 2006; 16(2):240–6. [PubMed: 16563739]

Grodzinsky Y, Santi A. The battle for Broca's region. Trends Cogn Sci. 2008; 12(12):474–80. [PubMed: 18930695]

Grodzinsky Y. The picture of the linguistic brain: How sharp can it be? Reply to Fedorenko & Kanwisher. Language and Linguistics Compass. 2010; 4(8):605–622. [PubMed: 20976129]

Hagoort P, Indefrey P, Brown C, Herzog H, Steinmetz H, Seitz RJ. The neural circuitry involved in the reading of german words and pseudowords: A PET study. J Cogn Neurosci. 1999; 11(4):383–98. [PubMed: 10471847]

Hagoort, P.; Baggio, G.; Willems, RM. The new cognitive neurosciences. IV. Cambridge, MA: MIT Press; 2009. Semantic unification.

Hagoort P, Hald L, Bastiaansen M, Petersson KM. Integration of word meaning and world knowledge in language comprehension. Science. 2004; 304(5669):438–41. [PubMed: 15031438]

Hasson U, Nusbaum HC, Small SL. Repetition suppression for spoken sentences and the effect of task demands. J Cogn Neurosci. 2006; 18(12):2013–29. [PubMed: 17129188]

Heim S, Alter K, Ischebeck AK, Amunts K, Eickhoff SB, Mohlberg H, et al. The role of the left Brodmann's areas 44 and 45 in reading words and pseudowords. Cogn Brain Res. 2005; 25(3): 982–93.

Heim S, Eickhoff SB, Amunts K. Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. Neuroimage. 2009; 48(3):616–624. [PubMed: 19560543]

Hein G, Knight RT. Superior temporal sulcus--it's my area: Or is it? J Cogn Neurosci. 2008; 20(12): 2125–36. [PubMed: 18457502]

Hickok G, Okada K, Serences J. Area Spt in the human planum temporale supports sensory-motor interaction for speech processing. Journal of Neurophysiology. 2009; 101:2725–2732. [PubMed: 19225172]

Holmes VM, O'Regan JK. Eye fixation patterns during the reading of relative-clause sentences. Journal of Verbal Learning and Verbal Behavior. 1981; 20(4):417–430.

Horwitz B, Amunts K, Bhattacharyya R, Patkin D, Jeffries K, Zilles K, et al. Activation of broca's area during the production of spoken and signed language: A combined cytoarchitectonic mapping and PET analysis. Neuropsychologia. 2003; 41(14):1868–76. [PubMed: 14572520]

Humphries C, Willard K, Buchsbaum B, Hickok G. Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. Neuroreport. 2001; 12(8):1749–1752. [PubMed: 11409752]

Humphries C, Love T, Swinney D, Hickok G. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. Human Brain Mapping. 2005; 26:128–138. [PubMed: 15895428]

Humphries C, Binder JR, Medler DA, Liebenthal E. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. J Cogn Neurosci. 2006; 18(4):665–79. [PubMed: 16768368]

Humphries C, Binder JR, Medler DA, Liebenthal E. Time course of semantic processes during sentence comprehension: An fmri study. Neuroimage. 2007; 36(3):924–32. [PubMed: 17500009]

Indefrey P, Brown CM, Hellwig F, Amunts K, Herzog H, Seitz RJ, et al. A neural correlate of syntactic encoding during speech production. Proc Natl Acad Sci U S A. 2001; 98(10):5933–6. [PubMed: 11331773]

Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. Cognition. 2004; 92 (1/2):101–104. [PubMed: 15037128]

Iwamura Y, Tanaka M, Sakamoto M, Hikosaka O. Converging patterns of finger representation and complex response properties of neurons in area 1 of the first somatosensory cortex of the conscious monkey. Experimental Brain Research. 1983; 51(3):327–337.

January D, Trueswell JC, Thompson-Schill SL. Co-Localization of stroop and syntactic ambiguity resolution in Broca's area: Implications for the neural basis of sentence processing. J Cogn Neurosci. 2009; 21(12):2434–44. [PubMed: 19199402]

Kaan E, Swaab TY. The neural circuitry of syntactic comprehension. Trends in Cognitive Sciences. 2002; 6(8):350–356. [PubMed: 12140086]

Kaminski J, Call J, Fischer J. Word learning in a domestic dog: Evidence for "fast mapping". Science. 2004; 304(5677):1682–3. [PubMed: 15192233]

Kanwisher N. Functional specificity in the human brain: A window into the functional architecture of the mind. Proc Natl Acad Sci USA. 2010; 107(25):11163–11170. [PubMed: 20484679]

Kanwisher N, McDermott J, Chun MM. The fusiform face area: A module in human extrastriate cortex specialized for face perception. J Neurosci. 1997; 17(11):4302–11. [PubMed: 9151747]

King J, Just M. Individual differences in syntactic processing: the role of working memory. Journal of Memory and Language. 1991; 30:580–602.

Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience: The dangers of double dipping. Nat Neurosci. 2009; 12(5):535–40. [PubMed: 19396166]

Lindenberg R, Fangerau H, Seitz RJ. "Broca's area" as a collective term? Brain Lang. 2007; 102(1): 22–9. [PubMed: 17257665]

Matelli M, Luppino G, Rizzolatti G. Architecture of superior and mesial area 6 and the adjacent cingulate cortex in the macaque monkey. J Comp Neurol. 1991; 311(4):445–62. [PubMed: 1757597]

Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, et al. The cortical representation of speech. J Cogn Neurosci. 1993; 5:467–79.

Meyer M, Friederici AD, von Cramon DY. Neurocognition of auditory sentence comprehension: event related fMRI reveals sensitivity to syntactic violations and task demands. Cognitive Brain Research. 2000; 9:19–33. [PubMed: 10666553]

Michael EB, Keller TA, Carpenter PA, Just MA. Fmri investigation of sentence comprehension by eye and by ear: Modality fingerprints on cognitive processes. Hum Brain Mapp. 2001; 13(4):239–52. [PubMed: 11410952]

Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Annual Review of Neuroscience. 2001; 24:167–202.

Myers EB, Blumstein SE, Walsh E, Eliassen J. Inferior frontal regions underlie the perception of phonetic category invariance. Psychol Sci. 2009; 20(7):895–903. [PubMed: 19515116]

Neville HJ, Bavelier D, Corina D, Rauschecker J, Karni A, Lalwani A, et al. Cerebral organization for language in deaf and hearing subjects: Biological constraints and effects of experience. Proc Natl Acad Sci U S A. 1998; 95(3):922–9. [PubMed: 9448260]

Noppeney U, Phillips J, Price C. The neural areas that control the retrieval and selection of semantics. Neuropsychologia. 2004; 42:1269–1280. [PubMed: 15178178]

Noppeney U, Price CJ. An FMRI study of syntactic adaptation. J Cogn Neurosci. 2004; 16(4):702–13. [PubMed: 15165357]

Novick JM, Trueswell JC, Thompson-Schill SL. Cognitive control and parsing: Reexamining the role of broca's area in sentence comprehension. Cogn Affect Behav Neurosci. 2005; 5(3):263–81. [PubMed: 16396089]

Olson IR, Plotzker A, Ezzyat Y. The enigmatic temporal pole: A review of findings on social and emotional processing. Brain. 2007; 130(Pt 7):1718–31. [PubMed: 17392317]

Osterhout L, Holcomb PJ. Event-related potentials and syntactic anomaly: Evidence of anomaly detection during the perception of continuous speech. Language and Cognitive Processes. 1993; 8:413–438.

Osterhout L, Allen M, McLaughlin J, Inoue K. Brain potentials elicited by prose-embedded linguistic anomalies. Memory and Cognition. 2002; 30(8):1304–1312.

Ouattara K, Lemasson A, Zuberbühler K. Campbell's monkeys use affixation to alter call meaning. Plos One. 2009; 4(11):e7808. [PubMed: 19915663]

Parzen E. On estimation of a probability density function and mode. Ann Math Statist. 1962; 33(3): 1065–1076.

Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. Nat Rev Neurosci. 2007; 8(12):976–87. [PubMed: 18026167]

Pepperberg, I. The Alex studies: Cognitive and communicative abilities of grey parrots. Harvard University Press; Cambridge, MA: 2000.

Petersen SE, Fox PT, Snyder AZ, Raichle ME. Activation of extrastriate and frontal cortical areas by visual words and word-like stimuli. Science. 1990; 249(4972):1041–4. [PubMed: 2396097]

Plante E, Creusere M, Sabin C. Dissociating sentential prosody from sentence processing: activation interacts with task demands. NeuroImage. 2002; 17:401–410. [PubMed: 12482093]

Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage. 1999; 10(1):15–35. [PubMed: 10385578]

Poldrack RA. Can cognitive processes be inferred from neuroimaging data? Trends in Cognitive Sciences. 2006; 10(2):59–63. [PubMed: 16406760]

Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, et al. Probabilistic mapping and volume measurement of human primary auditory cortex. Neuroimage. 2001; 13:669–683. [PubMed: 11305896]

Rodd JM, Davis MH, Johnsrude IS. The neural mechanisms of speech comprehension: Fmri studies of semantic ambiguity. Cereb Cortex. 2005; 15(8):1261–9. [PubMed: 15635062]

Rogalsky C, Hickok G. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. Cerebral Cortex. 2009; 19:786–796. [PubMed: 18669589]

Rozzi S, Ferrari PF, Bonini L, Rizzolatti G, Fogassi L. Functional organization of inferior parietal lobule convexity in the macaque monkey: Electrophysiological characterization of motor, sensory and mirror responses and their correlation with cytoarchitectonic areas. Eur J Neurosci. 2008; 28(8):1569–88. [PubMed: 18691325]

Santi A, Grodzinsky Y. Working memory and syntax interact in broca's area. Neuroimage. 2007; 37(1):8–17. [PubMed: 17560794]

Savage-Rumbaugh ES, Murphy J, Sevcik RA, Brakke KE, Williams SL, Rumbaugh DM, et al. Language comprehension in ape and child. Monographs of the Society for Research in Child Development. 1993:101–252.

Saxe R, Kanwisher N. People thinking about thinking people. The role of the temporo-parietal junction in "theory of mind". Neuroimage. 2003; 19(4):1835–42. [PubMed: 12948738]

Saxe R, Brett M, Kanwisher N. Divide and conquer: A defense of functional localizers. Neuroimage. 2006; 30(4):1088–96. discussion 1097–9. [PubMed: 16635578]

Schlesewsky M, Bornkessel I. On incremental interpretation: Degrees of meaning accessed during sentence comprehension. Lingua. 2004; 114(9–10):1213–1234.

Schnur TT, Schwartz MF, Kimberg DY, Hirshorn E, Coslett HB, Thompson-Schill SL. Localizing interference during naming: Convergent neuroimaging and neuropsychological evidence for the function of broca's area. Proc Natl Acad Sci U S A. 2009; 106(1):322–7. [PubMed: 19118194]

Schormann T, Zilles K. Three-Dimensional linear and nonlinear transformations: An integration of light microscopical and MRI data. Hum Brain Mapp. 1998; 6(5–6):339–47. [PubMed: 9788070]

Snyder HR, Feigenson K, Thompson-Schill SL. Prefrontal cortical response to conflict during semantic and phonological tasks. J Cogn Neurosci. 2007; 19(5):761–75. [PubMed: 17488203]

Stromswold K, Caplan D, Alpert N, Rauch S. Localization of syntactic comprehension by positron emission tomography. Brain Lang. 1996; 52:452–473. [PubMed: 8653390]

Thompson-Schill SL. Dissecting the language organ: A new look at the role of broca's area in language processing. Twenty-First Century Psycholinguistics: Four Cornerstones. 2005:173–189.

Tootell RBH, Reppas JB, Dale AM, Look RB, Sereno MI, Brady TJ, Rosen BR. Functional MRI evidence for a visual motion after effect in human cortical area MT/V5. Nature. 1995; 375:139–141. [PubMed: 7753168]

Vandenberghe R, Nobre AC, Price CJ. The response of left temporal cortex to sentences. J Cogn Neurosci. 2002; 14(4):550–560. [PubMed: 12126497]

Vigneau M, Beaucousin V, Hervé PY, Duffau H, Crivello F, Houdé O, et al. Meta-Analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. Neuroimage. 2006; 30(4):1414–32. [PubMed: 16413796]

Vul, E.; Kanwisher, N. Begging the question: The non-independence error in fmri data analysis. In: Hanson, S.; Bunzl, M., editors. Foundations and Philosophy for Neuroimaging. 2009. p. 71-91.

Westbury C, Zatorre RJ, Evans AC. Quantifying variability in the planum temporale: a probability map. Cerebral Cortex. 1999; 9:392–405. [PubMed: 10426418]

Willems RM, Hagoort P. Broca's region: Battles are not won by ignoring half of the facts. Trends Cogn Sci. 2009; 13(3):101. author reply 102. [PubMed: 19223227]

Wright P, Randall B, Marslen-Wilson WD, Tyler L. Dissociating linguistic and task-related activity in the left inferior frontal gyrus. Journal of Cognitive Neuroscience. 2010; 23(2):404–413. [PubMed: 20201631]

Zilles K, Schleicher A, Langemann C, Amunts K, Morosan P, Palomero-Gallagher N, et al. Quantitative analysis of sulci in the human cerebral cortex: Development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture. Hum Brain Mapp. 1997; 5(4):218–221. [PubMed: 20408218]

Zilles K, Schleicher A, Palomero-Gallagher N, Amunts K. Quantitative analysis of cyto-and receptor architecture of the human brain. Brain Mapping: The Methods. 2002:2.

Appendix A. The relationship between sample size and selectivity

In the presence of inter-subject variability in the loci of activation, effect B might be present within a group-level ROI for a small proportion of subjects (due to "leakage" from nearby regions where effect B could be present). In these cases *selectivity* takes the general form:

$$S = \beta_A \cdot (1 - \beta_B) = T(\delta_A \cdot n^{1/2} + T^{-1}(\alpha)) \cdot (1 - T(\delta_B \cdot n^{1/2} + T^{-1}(\alpha))) \quad \text{Equation 2}$$

where β_B represents the sensitivity (power) to detect the effect B at the same false positive level α, δ_A and δ_B are the signal-to-noise ratios of the effects A and B, respectively, within the ROI, and *T* represents the cumulative distribution function of the t- statistic. Equation 2 reduces to Equation 1 (see section *Selectivity*) when effect B is not present within the ROI voxels (when considering $\delta_B = 0$, leading to $\beta_B = \alpha_B$). Even when effect B is present in a small proportion of subjects its effect on selectivity can be quite marked (see Figure A1), and it typically leads to decreasing levels of selectivity with increasing number of subjects (because the group-level analyses will now find a population-level significant effect B at voxels where only a small proportion of subjects truly show this effect).

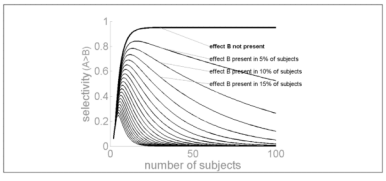


Figure A1.

Reduced selectivity with increasing number of subjects in the presence of inter-subject variability in the loci of activation. In this example we computed the selectivity (Equation 2) to detect an effect A (with δ_A =1, and α=.05) against a second effect B, when the effect B is not present (thick top line), as well as when the effect B is present only in a small proportion of subjects (thin lines) due to "leakage" from nearby areas in the presence of inter-subject variability in the loci of activation.

Appendix B. Functional group-level partitions

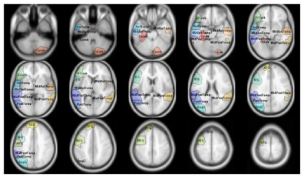
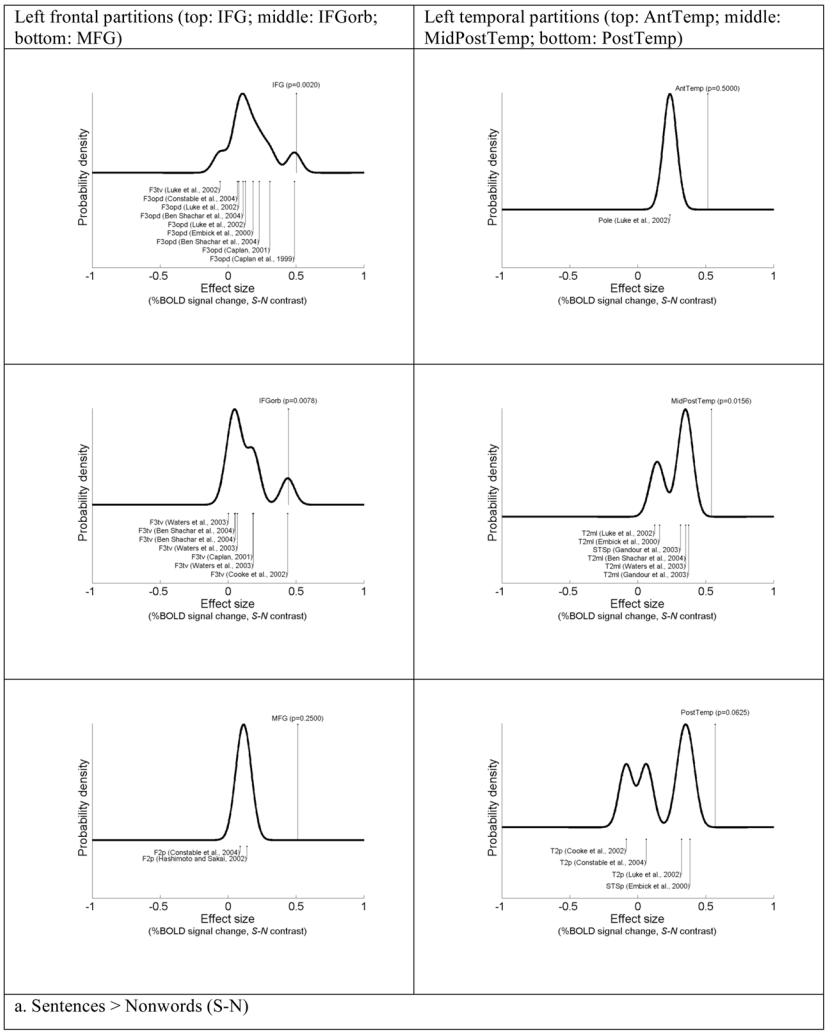


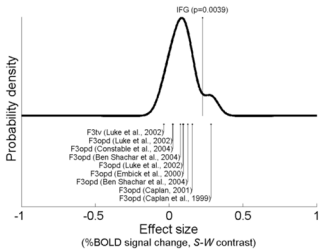
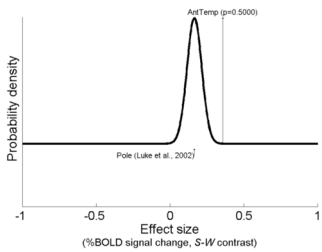
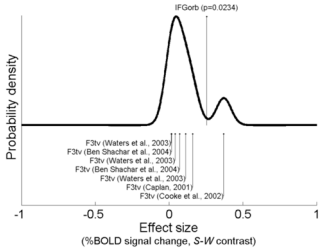
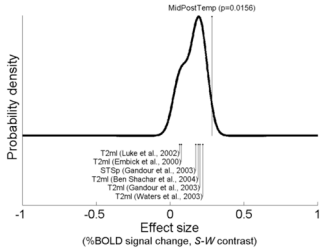
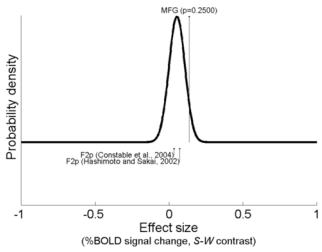
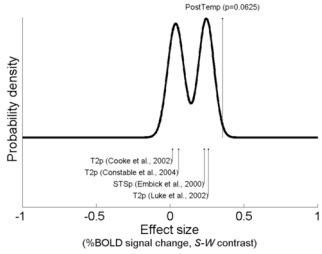
Figure B1.
Functional group-level partitions that are used to constrain the selection of subject-specific voxels in the GSS fROI method.

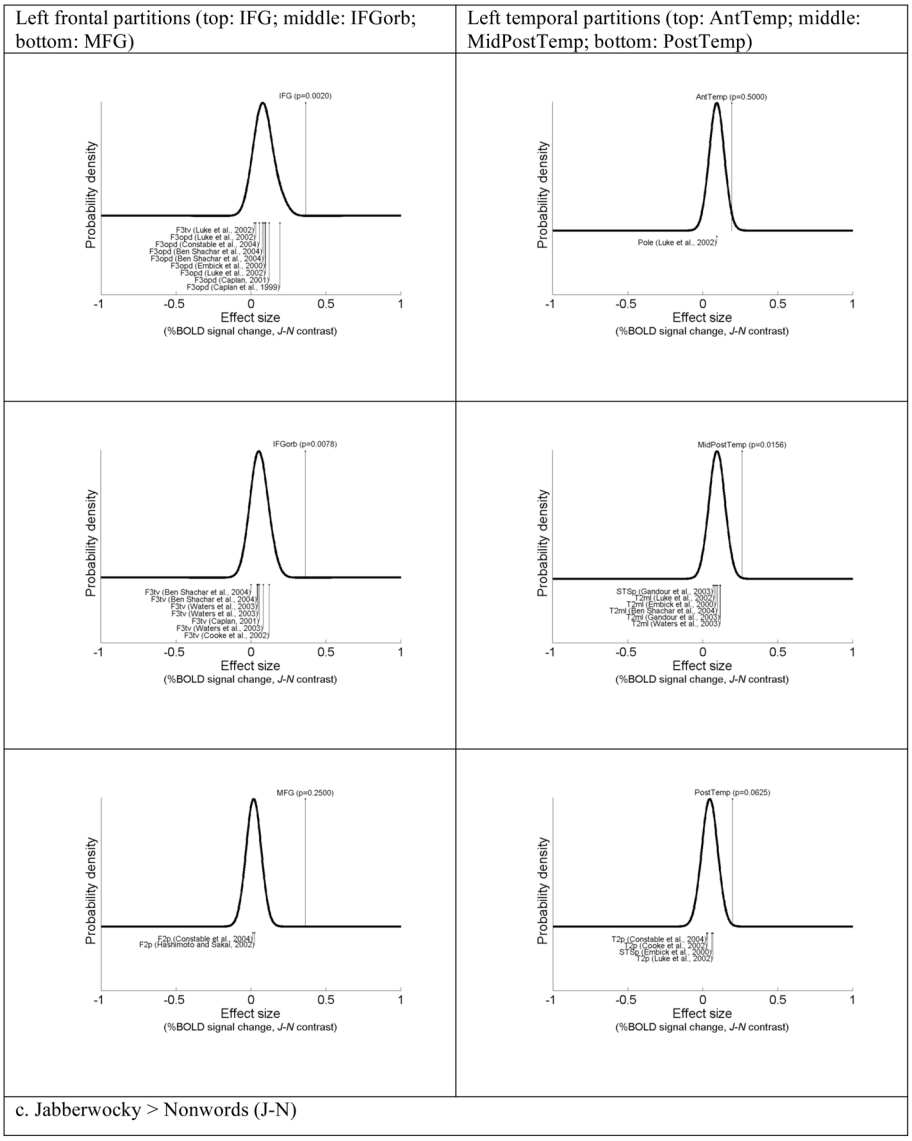
Appendix C. Effect size comparisons between GSS fROIs and the ROIs defined around the syntactic activation peaks from a number of previously

published studies that fell within the borders of the corresponding group-level partitions



a. Sentences > Nonwords (S-N)

Left frontal partitions (top: IFG; middle: IFGorb; bottom: MFG)	Left temporal partitions (top: AntTemp; middle: MidPostTemp; bottom: PostTemp)
	
	
	
b. Sentences > Words (S-W)	



c. Jabberwocky > Nonwords (J-N)

Figure C1.

Effect size comparisons between GSS fROIs and the ROIs defined around the syntactic activation peaks from a number of previously published studies that fell within the borders of the corresponding group-level partitions. The effect sizes for peak-defined ROIs are shown with arrows in the lower part of each plot. For each peak-defined ROI we provide information on the study where the peak was reported, and, for completeness, the cluster to which the peak was assigned in Vigneau et al.'s (2006) clustering analysis (e.g., F2p; see Vigneau et al., for details). For ease of visualization of the differences between average effect sizes for peak-defined ROIs and GSS fROIs for each partition and for each contrast, smoothed probability density plots were generated using kernel density estimation (Parzen, 1962) with bandwidth .05 %BOLD signal change (the bandwidth parameter is comparable to the "bin-width" parameter in the more common histogram plots and controls the amount of smoothing). The effect sizes for the GSS fROIs are shown with the arrow in the upper part of each plot.

Appendix D. An initial exploration of factors that plausibly affect variability in selectivity among the different language-sensitive brain regions

We here consider two factors that plausibly affect the degree of selectivity of our language-sensitive regions: (1) functional stability of a region, and (2) size of a region. With respect to both factors, we focus on our main contrast: sentences > nonwords. Also, although we are mostly interested in the variability in selectivity among the GSS fROIs, we consider group-level fROIs (investigated in Section B) because some factors may differentially affect selectivity for these two kinds of fROIs. In this way, we expect to dissociate differences in selectivity among fROIs from differences in selectivity among methods.

Functional stability

To estimate functional stability within a region, we computed correlation values across all voxels in each functional partition between odd and even runs (for the sentences minus nonwords contrast) (Fedorenko et al., 2010). For each region, we averaged these values across subjects. Functional stability in the patterns of differential response within a region can be expected to relate to the degree of selectivity within this region. In particular, if the selectivity of a region is low, the patterns of differential response can be expected to be weaker and also show reduced functional stability (either because of true differences in the loci of differential responses, or because of higher relative contributions of noise sources compared to these weaker differential effects). Furthermore, regarding the selectivity of the different methods, because the GSS method relies on the ability to consistently detect subject-specific loci of activation we are interested in determining whether the improvements in selectivity of the GSS method are limited to those areas showing high functional stability or extend to regions showing a wider range of functional stability values.

The associations between selectivity (indicated by S-N effect sizes) and functional stability (indicated by within-subject S-N spatial correlations) are shown in Figure D1. Functional stability is a good predictor of the degree of selectivity among the different regions, when considering both GSS fROIs ($r=.85$, $T(10)=5.11$, $p<.001$), as well as group-level fROIs ($r=.75$, $T(10)=3.61$, $p=.005$). This finding corroborates our expectation that the functional stability of an area is strongly associated with its selectivity.

When considering the selectivity of the different methods, the improvements in selectivity when using GSS fROIs compared to group-level ROIs, while modulated by functional stability (ANCOVA effect of *functional stability*, $T(10)=2.45$, $p=.034$), extend to the entire range of observed functional stability values (ANCOVA main effect of *method* controlled at the level of the minimum observed functional stability value (.24), $T(10)=4.52$, $p=.001$). This finding is interesting because it indicates that, although the comparative improvements in selectivity of the GSS fROI method are *more marked* fort hose areas showing high levels of functional stability, the improvements are manifested for a wide range of levels of functional stability. In particular, if there is a minimal level of functional stability necessary for the GSS fROI method to outperform the group-level fROI method, this minimal level is well below the range of functional stability values observed in the current study (minimal level below .09, $p<.05$).

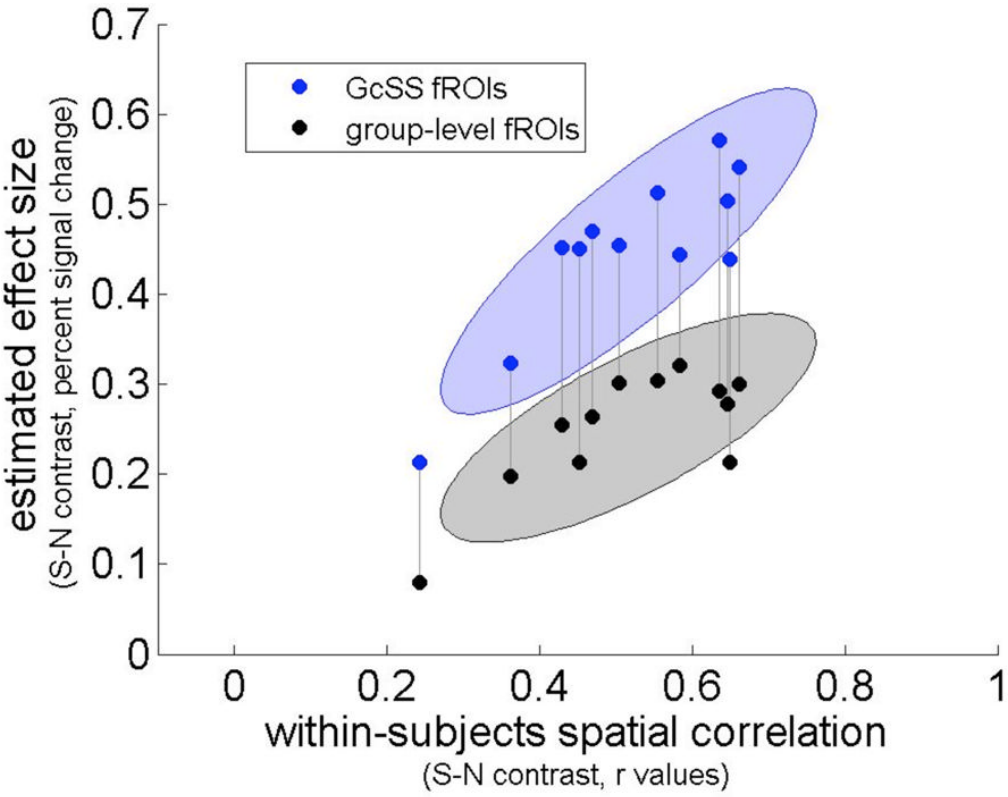


Figure D1.
Functional stability (estimated using within-subject spatial correlations) as a predictor of variability in the selectivity of the different language-sensitive regions.

Size

When considering a set of regions of different sizes around a fixed anatomical location, Poldrack (2006) found an inverse association between ROI size and selectivity. To our knowledge, it is not known whether this association extends to comparisons among different areas. Investigations of possible associations between selectivity and ROI size can guide the definition of appropriate units of analyses that would offer maximal selectivity. Regarding the selectivity of the different methods, because the GSS method tends to produce on average smaller ROIs in our sample than the group-level method we are interested in determining if the improvements in selectivity of the GSS method are in some way related to the decrease in ROI size (when comparing the selectivity of the same regions).

The associations between selectivity and ROI size are shown in Figure D2. When considering the selectivity of different regions, ROI size was positively associated with selectivity for GSS fROIs ($r=.75$, $T(10)=3.61$, $p=.005$) but not for group-level fROIs ($r=.02$, $T(10)=0.05$, $p=.96$). This finding could indicate a true association between ROI size and selectivity (across the language-sensitive regions) that is revealed by the GSS fROI method. Alternatively, it could indicate a selectivity dependency of the GSS fROI method that would result in selectivity improvements only for large ROIs (although this possibility is not supported by the additional between-method comparisons below). Further analyses and replications would be needed to better characterize the nature of this association.

When considering the selectivity of the different methods, the improvements in selectivity when using GSS fROIs compared to group-level ROIs were only marginally associated with the corresponding decreases in ROI size (r=−.25, T(10)=−0.83, p=0.43). The associations were weakly negative indicating that while the improvements might be partially related to the associated decreases in ROI size (ROI size decreased on average from 670 voxels in group-level fROIs to 380 voxels in GSS fROIs), in agreement with Poldrack (2006), this effect does not fully account for the observed improvements in the GSS fROIs selectivity (only 6% of the variance in selectivity values is explained by the reduction in ROI sizes). In addition, improvements in selectivity were only weakly modulated by baseline group-level ROI size (ANCOVA effect of *ROI size*, T(10)=1.34, p=.21), or by target GSS ROI size (ANCOVA effect of *ROI size*, T(10)=1.63, p=.13), and they extended to the entire range of observed ROI size values (ANCOVA main effect of *method* controlled at the level of the minimum observed ROI size value (103 voxels), T(10)=7.12, p<.001). These results indicate that the GSS fROI method outperforms the group-level fROI method for the entire range of observed ROI sizes. In addition they indicate that the improvements in selectivity, when comparing the two methods, are only weakly modulated by ROI size. This suggests that the observed association between ROI size and selectivity for the GSS fROIs does not result from a possible better performance of this method for large ROIs (cf. we could expect better performance for small ROIs due to the relatively larger effect of inter-subject variability in the loci of activation in these cases), but rather could be a true association for these regions that is revealed by the GSS fROI method but not apparent when using group-level fROIs.

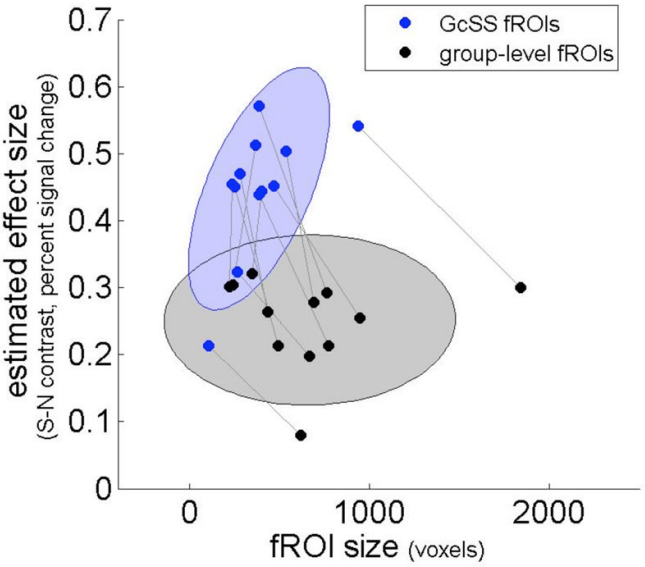


Figure D2.

Size as a predictor of variability in the selectivity of the different language-sensitive regions. Given how important selectivity of different brain regions is for understanding the neural architecture underlying human cognition, future work may be aimed at (a) investigating these factors further with respect to language-sensitive regions by examining other linguistic and non-linguistic contrasts, (b) investigating the generalizabilty of these factors by examining their effects on selectivity of regions in other domains, and (c) exploring other factors that may affect selectivity, such as location or spatial variability across subjects.

Appendix E. Quantifying the overlap between BAs 44/45 and our IFG fROIs

In order to quantify the amount of overlap between the two anatomical ROIs used in Section C (BA44, BA45), on the one hand, and two of our fROIs located in/around the left IFG (Left IFG, Left IFGorb), on the other hand, we computed the probability of each of the two fROIs falling within the borders of either BA44 or BA45, as described in Section C. For completeness, we computed these values for both (a) GSS fROIs, and (b) group-level fROIs (used in Section B). For GSS fROIs, the probability was computed for each subject's fROI and then the values were averaged. The results are presented in Table E1.

Table E1

Estimates of the probability that our functional ROIs (Left IFG and Left IFGorb) fall within BA44 vs. BA45.

fROI	*P* of being within:	BA44	BA45
Left IFG (GSS)		.19 (.01)	.16 (.02)
Left IFGorb (GSS)		.01 (.002)	.08 (.01)
Left IFG (group-level)		.23	.19
Left IFGorb (group-level)		.02	.12

Appendix F. Response profiles of BA 44 and BA 45 with respect to the four conditions

We here present the responses to the four conditions in our dataset (sentences, words, jabberwocky and nonwords) in the two anatomical ROIs used in Section C (BA 44 and BA 45). We performed the extraction using three different thresholds in the probabilistic maps for each BA: (a) all the voxels in each map, (b) only the voxels that are present in five out of ten brains, and (c) only the voxels that are present in nine out of ten brains (highest-overlap voxels). If either or both of these areas support syntactic processing, then they should show a higher response to sentences than to words, and a higher response to jabberwocky than to nonwords. The results are presented in Figure F1.

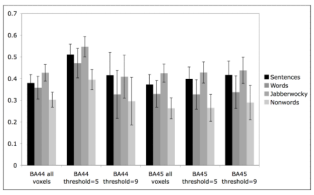


Figure F1.
Responses of BAs 44 and 45 to the four conditions in our dataset, using different threshold levels to define the ROIs. Error bars represent standard errors of the mean, computed across subjects.

For the jabberwocky > nonwords contrast, consistent with these regions' engagement in syntactic processing, we see significant effects in BA 45 at all three thresholds (p-values: <. 001, <.005, and <.01), and in BA44 at the two more liberal thresholds (p-values: <.001, <. 005, n.s.). However, for the sentences > words contrast there are no significant effects at any threshold for either BA11. Without a reliably higher response to sentences than to words, it is difficult to interpret the effect observed for the jabberwocky > nonwords contrast. Even the reliable effects for the former contrast are small, and overall, the conditions are not well differentiated.

The lack of a reliable difference between the sentences and the words conditions is surprising given that (1) this is a large contrast (cf. contrasts between syntactically complex and syntactically simpler sentences used in the previous studies), (2) the sample size is relatively large (n=25), and (3) these regions are among the prime candidates for syntactic processing. Based on these data, one would be hard pressed to argue for the role of these regions in syntactic processing, leave alone some narrow aspect of syntax (e.g., Grodzinsky & Santi, 2008). However, it is not the case that we think that these regions play no role in syntactic processing. Rather these results underscore, once again, the importance of using subject-specific functional ROIs instead of standard group ROIs that are the same across subjects. As reported in Section C, both of the GSS fROIs in the left IFG (as well as subject-specific fROIs defined by intersecting the BA maps with individual activation maps) strongly differentiate between the sentences and the words conditions. This suggests that the analyses reported here underestimate selectivity (plausibly for the reasons discussed throughout the paper), and are therefore not optimal for addressing questions of the functional architecture of the language system.

Appendix G. Standardized effect sizes for all the ROIs discussed in the paper

Table G1

Effect sizes for the GSS fROIs.

	fROI	Sentences > Nonwords	Sentences > Words	Jabberwocky > Nonwords
Left frontal regions	IFG	1.43	0.75	1.30
	IFGorb	1.81	0.81	1.22
	MFG	1.21	0.40	1.49
	SFG	1.05	0.76	0.69
Left posterior regions	AntTemp	1.50	1.20	1.00
	MidAntTemp	1.91	1.28	1.23
	MidPostTemp	1.92	1.20	1.45
	PostTemp	2.16	1.32	0.60
	AngG	1.29	0.55	0.69
Right posterior regions	MidAntTemp	1.38	1.41	0.63
	MidPostTemp	1.69	0.77	1.19
Cerebellar regions	Right Cereb	1.53	0.49	1.05
	Left Cereb	0.96	0.57	0.40

[11]It is worth noting that the fact that these BAs do not show dissociable functional profiles with respect to this set of conditions does not mean that it is not possible to obtain such a dissociation in principle. In fact, several published papers report dissociations between activations observed in BAs 44 vs. 45 (e.g., Binkofski et al, 2000; Horwitz et al., 2003; Heim et al., 2005; Santi & Grodzinsky, 2007; Heim et al., 2009).

Table G2

Effect sizes for the peak-defined ROIs (syntactic peaks from Vigneau et al., 2006). [Results section A]

	fROI	Sentences > Nonwords	Sentences > Words	Jabberwocky > Nonwords
ROIs for peaks within the IFG partition	Ben-Shachar et al., 2004 F3opd	0.30	0.41	0.41
	Ben Shachar et al., 2004 F3opd	0.66	0.56	0.55
	Caplan, 2001 F3opd	0.97	0.57	0.69
	Caplan et al., 1999 F3opd	1.22	0.85	0.74
	Constable et al., 2004 F3opd	0.34	0.18	0.38
	Embick et al., 2000 F3opd	0.60	0.49	0.46
	Luke et al., 2002 F3opd	0.36	0.43	0.46
	Luke et al., 2002 F3opd	0.36	0.12	0.24
	Luke et al., 2002 F3tv	−0.14	−0.10	0.10
IFGorb	Ben Shachar et al., 2004 F3tv	0.31	0.60	0.01
	Ben Shachar et al., 2004 F3tv	0.24	0.15	0.33
	Caplan, 2001 F3tv	0.68	0.78	0.27
	Cooke et al., 2002 F3tv	0.84	0.86	0.38
	Waters et al., 2003 F3tv	0.32	0.21	0.24
	Waters et al., 2003 F3tv	0.01	0.08	0.31
	Waters et al., 2003 F3tv	0.68	0.45	0.51
MFG	Constable et al., 2004 F2p	0.41	0.21	0.11
	Hashimoto and Sakai, 2002 F2p	0.43	0.32	0.21
AntTemp	Luke et al., 2002 Pole	1.23	0.84	0.60
MidPostTemp	Gandour et al., 2003 STSp	0.87	0.89	0.51
	Ben Shachar et al., 2004 T2ml	1.79	1.55	0.76
	Embick et al., 2000 T2ml	0.62	0.49	0.49
	Gandour et al., 2003 T2ml	1.53	1.16	0.77
	Luke et al., 2002 T2ml	0.52	0.34	0.44

	fROI	Sentences > Nonwords	Sentences > Words	Jabberwocky > Nonwords
	Waters et al., 2003 T2ml	1.61	1.37	0.64
PostTemp	Embick et al., 2000 STSp	0.88	0.96	0.31
	Constable et al., 2004 T2p	0.35	0.48	0.20
	Cooke et al., 2002 T2p	−0.31	0.11	0.17
	Luke et al., 2002 T2p	0.87	1.14	0.39

Table G3

Effect sizes for the group-level fROIs. [Results section B]

	fROI	Sentences > Nonwords	Sentences > Words	Jabberwocky > Nonwords
Left frontal regions	IFG	1.05	0.41	1.07
	IFGorb	1.10	0.70	1.05
	MFG	0.84	0.33	1.12
	SFG	0.79	0.56	0.64
Left posterior regions	AntTemp	1.01	0.76	0.67
	MidAntTemp	1.26	0.64	0.76
	MidPostTemp	2.05	0.86	1.69
	PostTemp	1.18	0.76	0.83
	AngG	0.94	0.57	0.38
Right posterior regions	MidAntTemp	1.07	0.83	0.45
	MidPostTemp	1.44	0.56	1.12
Cerebellar regions	Right Cereb	1.36	0.44	1.23
	Left Cereb	0.59	0.37	0.44

Table G4

Effect sizes for the anatomical group-level ROIs and for individually defined fROIs using the probabilistic cytoarchitectonic maps to constrain the selection of voxels in individual subjects. [Results section C]

	fROI	Sentences > Nonwords	Sentences > Words	Jabberwocky > Nonwords
Anatomically defined group-level ROIs	BA 44 all voxels	0.51	0.13	0.81
	BA 45 all voxels	0.55	0.22	0.84
	BA 44 high overlap voxels	0.53	0.30	0.56
	BA 45 high	0.35	0.30	0.32

	fROI	Sentences > Nonwords	Sentences > Words	Jabberwocky > Nonwords
	overlap voxels			
Subject-specific fROIs (BA maps are used to constrain the selection of individual voxels)	BA 44	1.44	0.73	1.25
	BA 45	1.59	0.80	1.17

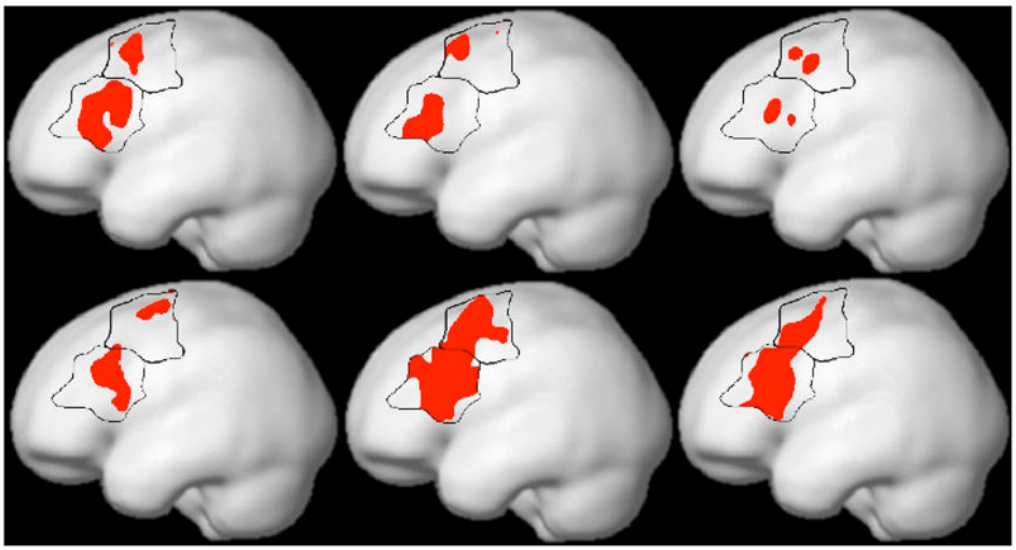


Figure 1.
Two sample fROIs (Left IFG and Left MFG) in six sample subjects. The borders of the group-level partitions are shown with a black outline and the subject-specific activations are shown in red.

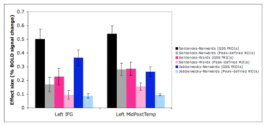
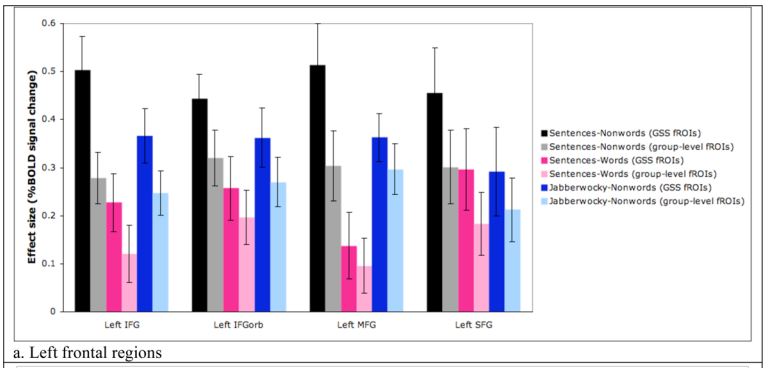
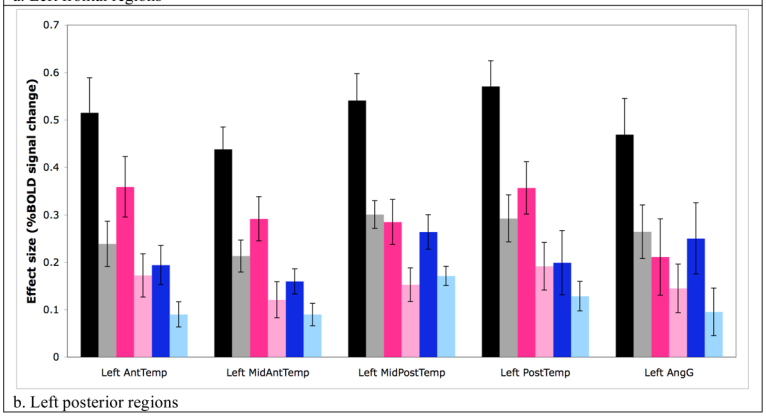


Figure 2.
Effect size comparisons between GSS fROIs (darker bars) and the ROIs defined around the syntactic activation peaks from several previously published studies that fell within the borders of the corresponding group-level partitions (lighter bars) for the sentences > nonwords contrast (black and grey bars), sentences > words contrast (pink bars), and jabberwocky > nonwords contrast (blue bars), for two sample regions (Left IFG and Left MidPostTemp).



a. Left frontal regions



b. Left posterior regions

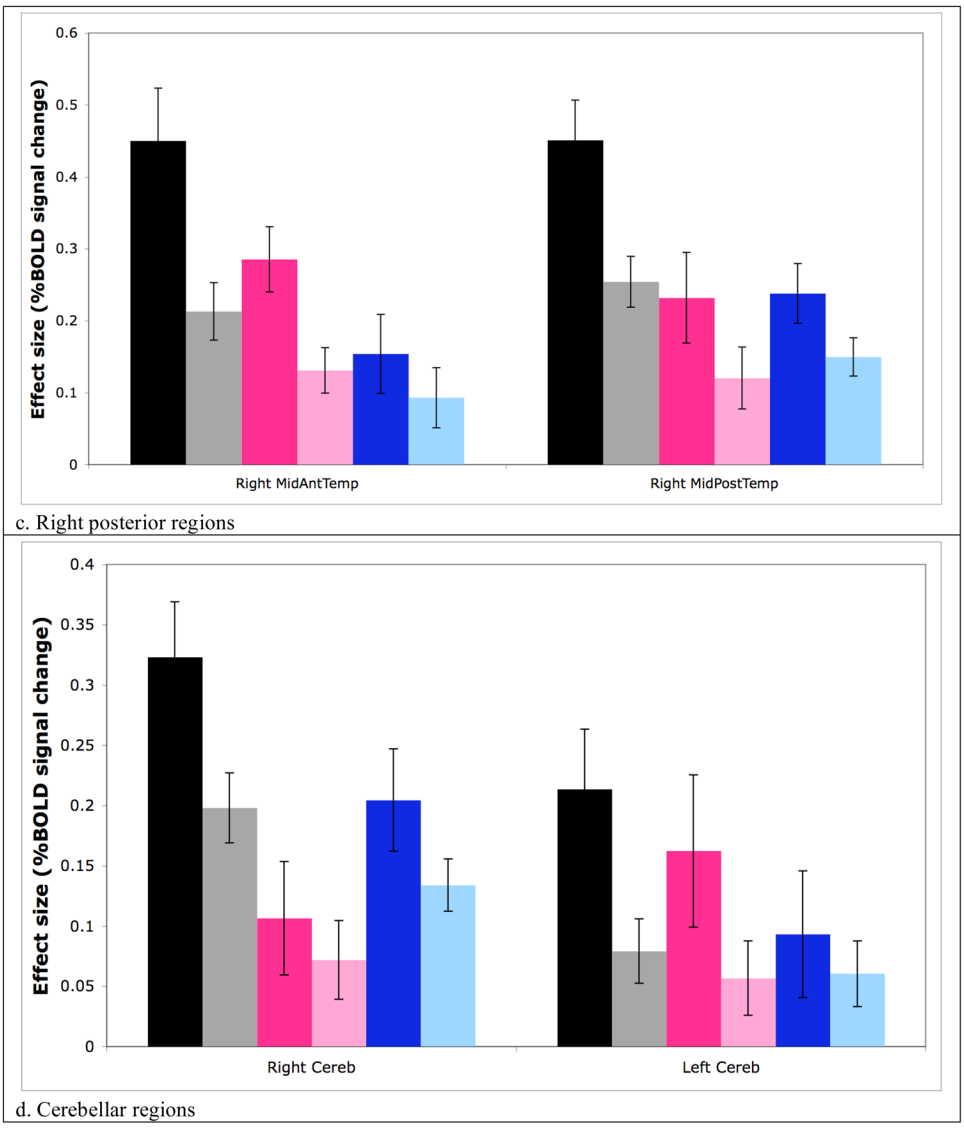


Figure 3.
Effect size comparisons between GSS fROIs (darker bars) vs. group fROIs (lighter bars) for the sentences > nonwords contrast (black and grey bars), sentences > words contrast (pink bars), and jabberwocky > nonwords contrast (blue bars).

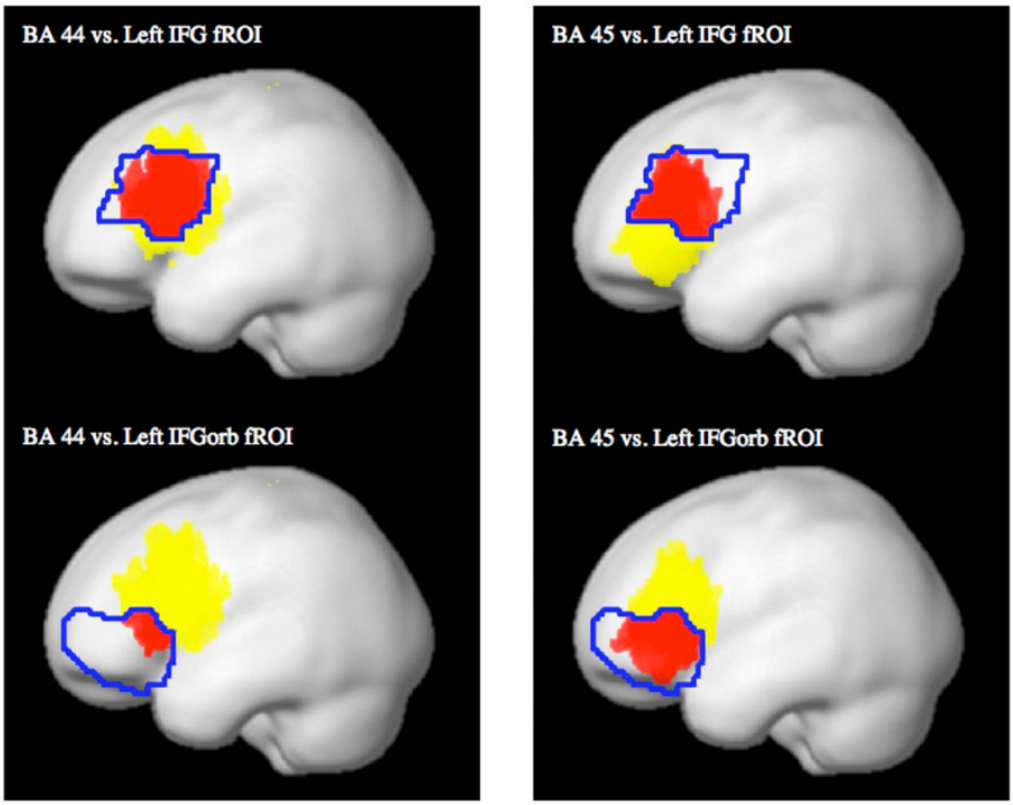


Figure 4.
The relationship between Amunts et al. (1999) probabilistic cytoarchitectonic maps (shown in yellow) for BA 44 (left panel) and BA 45 (right panel) and our group-level partitions (shown with blue contours) for Left IFG (top) and Left IFGorb (bottom). Overlap is shown in red.

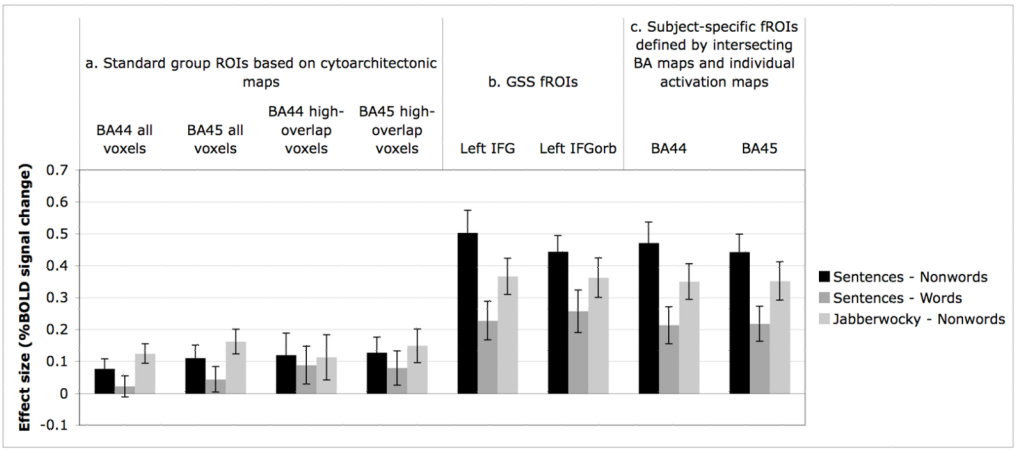


Figure 5.

Effect size comparison for BA44 and BA45 defined in two different ways using Amunts et al.'s (1999) probabilistic cytoarchitectonic maps (5a); two GSS fROIs (Left IFG, and Left IFGorb) (5b); and subject-specific fROIs created by intersecting the BA maps with individual activation maps (5c).

Table 1

Results of Wilcoxon signed rank tests comparing the effect size values between GSS fROIs and peak-defined ROIs located within the corresponding partition.

	fROI	Number of peaks within partition	Sentences > Nonwords	Sentences > Words	Jabberwocky > Nonwords
Left frontal regions	IFG	9	<.005	<.005	<.005
	IFGorb	7	<.01	<.05	<.01
	MFG	2	n.s.	n.s.	n.s.
Left posterior regions	AntTemp	1	n.s.	n.s.	n.s.
	MidPostTemp	6	<.05	<.05	<.05
	PostTemp	4	.06	.06	.06

Table 2

Results of the t-tests comparing the effect size values between GSS fROIs and group fROIs.

	fROI	Sentences > Nonwords	Sentences > Words	Jabberwocky > Nonwords
Left frontal regions	IFG	<.0005	<.001	<.0005
	IFGorb	<.05	n.s.	.055
	MFG	<.005	n.s.	n.s.
	SFG	<.005	<.05	<.05
Left posterior regions	AntTemp	<.0001	<.0005	<.01
	MidAntTemp	<.0001	<.0001	<.0005
	MidPostTemp	<.0001	<.0001	<.0005
	PostTemp	<.0001	<.0005	n.s.
	AngG	<.005	n.s.	<.05
Right posterior regions	MidAntTemp	<.001	<.0005	<.05
	MidPostTemp	<.0005	<.05	<.05
Cerebellar regions	Right Cereb	<.005	.090	.064
	Left Cereb	<.01	<.05	n.s.