

# **Prelexical and lexical processing in listening**

Peter Indefrey and Anne Cutler

*Max Planck Institute for Psycholinguistics,*

*Nijmegen, The Netherlands*

Running title: Prelexical and lexical processing

*E-mail addresses:*

[Peter.Indefrey@mpi.nl](mailto:Peter.Indefrey@mpi.nl), [Anne.Cutler@mpi.nl](mailto:Anne.Cutler@mpi.nl)

## **Abstract**

This paper presents a meta-analysis of hemodynamic studies on passive auditory language processing. We assess the overlap of hemodynamic activation areas and activation maxima reported in experiments involving the presentation of sentences, words, pseudowords, or sublexical or non-linguistic auditory stimuli. Areas that have been reliably replicated are identified. The results of the meta-analysis are compared to electrophysiological, magnetencephalic (MEG), and clinical findings. It is concluded that auditory language input is processed in a left posterior frontal and bilateral temporal cortical network. Within this network, no processing level is related to a single cortical area. The temporal lobes seem to differ with respect to their involvement in post-lexical processing, in that the left temporal lobe has greater involvement than the right, and also in the degree of anatomical specialization for phonological, lexical, and sentence-level processing, with greater overlap on the right contrasting with a higher degree of differentiation on the left.

## **1. Introduction**

The listener's task in understanding spoken language is to extract meaning from an almost continuous stream of sound. The meaning - the message which speakers wish to convey - is expressed via a sequence of words and the relationships between those words (which can in turn be encoded, depending on the language, in the order in which the words occur, or in inflectional markings of various kinds applied to them). Whole utterance meanings are rarely stored in memory; the beauty of language is the possibility of expressing new meanings via combinations of known words. The words are however stored as known units, in some form (again, the form may be language-dependent). Understanding spoken language thus requires recognition of the individual words which the speaker has uttered.

Laboratory tasks via which this type of processing has been investigated over the past four decades fall into two main groups: tasks which require attention to sublexical units, and tasks which require attention to words. The first category includes the detection of target phonemes or fragments (e.g. "press the button when you hear a word containing /b/" or "... containing bot-"), or the categorization of phonemes ("does this syllable begin with /d/ or /t/?"). The second category includes lexical decision ("is this a real word or not?"), word spotting ("press the button when you hear any real word in this input"), word reconstruction, mispronunciation detection and phoneme restoration (all tasks which require finding the word which most closely matches a slightly distorted input), and various adaptations or combinations of these (e.g. priming tasks, in which for instance a lexical decision on a second word may be affected by a preceding first word), as well as decisions about semantic properties of words ("abstract or concrete?"). All these tasks can involve measurement of response time; further, accuracy rates may also be informative, and in phoneme categorization the nature of the decision is crucial. For an overview of the tasks see Grosjean and Frauenfelder (1997).

Via such research methods, an enormous amount of knowledge has been garnered which can guide neuroimaging approaches to prelexical and lexical processing in listening. Briefly summarized, research on spoken-word recognition has produced no agreement on the nature of the prelexical representations (if any)

involved in listening, but almost unanimous agreement on the following premises: (1) speech information is continuously mapped to the lexicon; (2) speech input can activate multiple candidate words simultaneously, including partial activation of words with partial support from the input; (3) concurrently active word candidates compete with one another for recognition. Continuity is supported by evidence that coarticulatory information for upcoming phonemes can constrain lexical activation (e.g. Streeter and Nigro, 1979; Whalen, 1991; Marslen-Wilson and Warren, 1994; McQueen, Norris, and Cutler, 1999); multiple activation is supported by evidence that word fragments activate multiple possible completions (e.g. Zwitserlood, 1989; Connine, Blasko, and Wang, 1994; Zwitserlood and Schriefers, 1995; Soto-Faraco, Sebastian-Galles, and Cutler, 2001) and by evidence of activation of words only spuriously present in the speech input, such as lips in tulips (Gow and Gordon, 1995; Tabossi, Burani, and Scott, 1995; Vroomen and De Gelder, 1997; Luce and Lyons, 1999); and competition is supported by evidence that the more words are potentially compatible with the input, the harder recognition of any one of them becomes (McQueen, Norris, and Cutler, 1994; Norris, McQueen, and Cutler, 1995; Vroomen and De Gelder, 1995; Soto-Faraco et al., 2001). A recent review of the evidence is presented by McQueen, Dahan, and Cutler (2003).

Agreement on the nature of prelexical representations is lacking because there is evidence compatible (a) with continuous cascade of information through intervening levels at which, for instance, discrete phonemic representations play a role, (b) with alternative abstract representations such as syllables or featural bundles, or (c) with a model involving no intervening representations, where lexical mapping is achieved via computation of similarity to previously encountered acoustic forms.

Of course, behavioral tasks do not tap directly into the necessary processing steps involved in speech recognition; they all involve some decision component, so that observed effects could arise at the decision level rather than at levels normally involved in recognition. Thus lexical effects in phoneme-level tasks need not imply that the prelexical evaluation of speech signals is directly influenced by lexical knowledge; it could be that explicit decisions about phonemic identity are informed by lexical information (see Norris, McQueen, and Cutler, 2000, for further

discussion). It is for this reason necessary, and indeed standard, in behavioral research to seek converging evidence from a wide variety of tasks with differing behavioral profiles, and thanks to this strategy these tasks have provided a wealth of evidence concerning lexical processing.

Neurophysiological and neuroimaging techniques attempt to achieve an unmediated reflection of processing. In this respect reliance on the behavioral tasks normally used in spoken-word recognition research introduces potentially serious complications. In a previous contribution to this series, Norris and Wise (2000) pointed out that the problem-solving and decision aspects of the most widely-used tasks in this area of cognitive neuroscience, such as phoneme detection, lexical decision or semantic categorization, may recruit many cognitive subsystems beyond those involved in speech processing, and that observed contrasts may reflect differences in secondary tasks rather than in relevant underlying processes. Giraud and Price (2001) made similar cautionary remarks.

In the present review, therefore, we endeavor to summarize what can be learned about prelexical and lexical processing from neuropsychological, neurophysiological and neuroimaging studies in the absence of task confounds. To this purpose we assembled all available evidence uncontaminated by potential artifacts due to secondary task, as we describe below. We extended the scope of this evidence beyond studies involving just sublexical and lexical stimuli, by including tone stimuli (to assess the involvement of simple auditory processing) and sentence stimuli (to assess the linguistic processing of which prelexical and lexical processing form a part).

## **2. Cerebral regions involved in language perception – a meta-analysis**

### **2.1. Procedures**

*Data set.* We queried the Current Contents and Medline publication databases using the following combination of search terms: (PET OR Positron Emission Tomography OR fMRI OR functional Magnetic Resonance Imaging OR MEG OR Magnetencephalography OR magnetencephalographic) AND (auditory OR voice OR

language *OR* sound *OR* speech). This query resulted in a set of 1,058 neuroimaging studies using linguistic and non-linguistic auditory stimuli (December 19, 2002). Based on the abstracts, 463 potentially relevant publications were selected. Among these, 36 publications (marked by \* in the reference list) presenting 55 experiments met the following criteria: (1) there was no task other than listening to the auditory stimuli, (2) the activations were reported in Talairach coordinates or an equivalent coordinate system. In more than half of these experiments (see Table 1), auditory stimuli were compared to a condition in which no acoustic stimuli other than the inevitable background noise of the PET or fMRI scanning procedure were presented (henceforth called 'silent' control condition). Given this constant control condition, we assumed that activations from similar stimuli should similarly reflect acoustic and linguistic properties of the stimuli. Using a reliability estimate, we assessed the degree of anatomical overlap of activation areas in four experiments presenting sentences, ten experiments presenting words, four experiments presenting pseudowords or meaningless syllables (monosyllabic pseudowords), and ten experiments presenting non-linguistic tone stimuli. Note that a silent rest baseline is not without problems, since some brain regions seem to be activated during such a baseline. Binder et al. (1999) examined hemodynamic activations of silent rest compared to a simple perceptual task (listening to tones), and found significant blood flow increases in the left posterior frontal lobe, the left angular gyrus, and the bilateral cingulate gyri. This means that activations of auditory stimuli in these areas may be at least in part obscured when compared to silent rest. Binder et al. (1999) report that silent rest does not seem to activate the temporal lobes more strongly than tone stimuli do. Shulman et al. (1997) found greater temporal activation in left BA 20 for a silent fixation condition compared to passive viewing of non-linguistic visual stimuli, but not compared to passive viewing of linguistic stimuli. Thus although there may be some temporal lobe activation during silent rest, it is too weak to obscure hemodynamic responses during passive listening.

**Table 1.** Overview of passive listening experiments with silent control conditions that contributed to the meta-analyses (Table 3 and Figure 1) and the localization synopsis in Figure 2. All experiments except numbers 3, 9, 12, 17, 31, and 34 were entered in the reliability analyses. Data from experiments 1 and 2 as well as 19 and 20 were collapsed in the reliability analyses, since the experiments were not independent.

<i>Study</i>	<i>Stimulus</i>	<i># in Figure 2</i>
Belin 1998	200ms frequency transition, 60/min	1
Belin 1998	40ms frequency transition, 60/min	2
Belin 1999	synthetic diphthong, 6/min	3
Binder 2000	tones, different frequencies, 90/min	4
Bookheimer 1998	pseudowords, 9/min	5
Celsis 1999	syllables, 180/min	6
Celsis 1999	tones, 500 + 700Hz, 180/min	7
di Salle 2001	tones, 1000Hz, 6/min	8
Engelien 1995	environmental sounds, 10/min	9
Fiez 1996	pseudowords, 60/min	10
Fiez 1996	words, 60/min	11
Giraud 2000	vowels vs. expecting vowels, 120/min	12
Holcomb 1998	tones, 1500Hz + lower tones, 30/min	13
Jäncke 1999	tones, 1000Hz, 60/min	14
Lockwood 1999	tones, 500 + 4000Hz, 60/min	15
Mellet 1996	words, 30/min	16
Mirz 1999	music	17
Mirz 1999	sentences	18
Mirz 1999	tones, 1000Hz	19
Mirz 1999	tones, 1000 + 4000Hz	20
Mirz 1999	words	21
Müller 1997	sentences, 12/min	22
Petersen 1988	words, 60/min	23
Price 1996	words, 40/min	24
Price 1996	words, different rates	25
Suzuki 2002a	words, 60/min	26
Suzuki 2002b	tones, 1000Hz, 60/min	27
Thivard 2000	tones with spectral maxima, 60/min	28
Warburton 1996	words, 4/min	29
Wise 1991	pseudowords, 40 or 60/min	30
Wong 1999	reversed sentences, 30/min	31
Wong 1999	sentences, 30/min	32
Wong 1999	words, 30/min	33
Wong 2002	reversed words, 15/min	34
Wong 2002	sentences, 12/min	35
Wong 2002	words, 15/min	36

The remaining studies (see Table 2) investigated the neural substrate of specific aspects of the auditory stimuli presented in the active condition in comparison to various auditory control stimuli. Since in this subset no two studies used the same combination of active and control stimuli, we refrained from testing anatomical overlap between experiments, and present the activation foci reported in these studies descriptively.

**Table 2.** Overview of passive listening experiments with heterogeneous auditory control conditions that contributed to the localization synopsis in Figure 3.

<i>Study</i>	<i>Stimulus vs. control stimulus</i>	<i># in Figure 3</i>
Benson 2001	CVC > CV > V	1
Binder 1996	words vs. tones	2
Binder 2000	pseudo vs. tones	3
Binder 2000	reversed words vs. tones	4
Binder 2000	words vs. tones	5
Giraud 2000	amplitude modulated noise vs. noise	6
Giraud 2000	sentences vs. vowels	7
Giraud 2000	words vs. vowels	8
Hall 2002	frequency modulated vs. static tone	9
Hall 2002	harmonic vs. single tone	10
Jäncke 2002	syllables vs. 350 ms white noise bursts	11
Jäncke 2002	syllables vs. steady state portion of vowel	12
Jäncke 2002	syllables vs. tones	13
Müller 2002	90% 1000Hz + 10% 500Hz vs. 1000Hz	14
Mummery 1999	words vs. signal correlated noise	15
Price 1996	words vs. reversed words	16
Schlosser 1998	sentences vs. unknown language	17
Scott 2000	sentences vs. rotated sentences	18
Thivard 2000	frequency transition vs. stationary tone	19

*Anatomical coding.* The reported activation foci were coded in two ways to account for global activation patterns on a whole-brain level as well as for finer-grained anatomical differences within the temporal lobe. At the whole-brain level, the reported foci were entered in an anatomical reference system of 92 cortical regions based on the parcellation of the cerebral cortex described by Rademacher et al. (1992), plus 16 subcortical and cerebellar regions. At the fine-grained level, the foci were entered in a coordinate system with millimeter resolution covering the temporal

lobes. Activation foci reported in MNI-coordinates were converted to the Talairach and Tournoux (1988) space using the nonlinear algorithm of Brett (1999, available at [www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html](http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html)).

*Reliability estimate.* Reliability estimates were calculated using the procedure of Indefrey and Levelt (2000, 2003). For any subset of experiments to be compared at the whole-brain level, we divided the average number of activated regions per experiment by the number of regions to obtain the chance probability for any particular region to be reported in one experiment. Assuming this probability, the chance level for a region to be reported as activated in a number of experiments is given by a binomial distribution. If this level was below 5%, we rejected the possibility that the agreement of reports about a given region was due to chance. Note that the studies were not assigned weights reflecting design or number of subjects, so that activation overlaps which are reliable according to our criterion cannot necessarily be interpreted as statistically significant.

At the fine-grained level, we applied a similar reliability criterion to the reported locations of temporal lobe activation maxima on the lateral temporal planes. Assuming a two-dimensional cortical layer, we ignored location differences in the medial-lateral dimension, except for the most medial (absolute value of the x-coordinate < 40) activation foci located on the superior temporal plane, which were analysed separately. Due to anatomical variability between subjects, the location of activation maxima may shift by several millimetres to over a centimetre between studies even if identical or highly similar experimental paradigms are used (Stromswold et al., 1996; Caplan et al., 1998, 2000; Indefrey et al., 2003). To account for this variability, all reported activation foci were converted to focal activation areas extending plus or minus five millimetres from the original coordinates in the dorsal-ventral and the rostral-caudal dimensions. Dividing the mean focal activation area per experiment by the approximate total area of the lateral temporal plane,  $3500 \text{ mm}^2$ , we obtained an estimate for the chance probability of every  $1\text{-mm}^2$  pixel to be activated in a single experiment. Note that for smaller sets of experiments the statistical power decreases, such that a relatively larger number of positive reports is required for an

above-chance decision. The procedure controlled for the fact that the average number of activation foci may differ across stimuli, increasing the chances of coincidental agreements of findings between studies.

The reliability criterion we applied does not entail that atypical findings of activations in any single study are necessarily due to chance. The number of experiments *not* reporting activations was insufficient to consider a region as inactive at the chosen error probability level. Isolated observations therefore do not exclude the possibility that a region is active. They may, for example, reflect smaller activations only detectable with refined techniques or better scanning devices. Furthermore, the nature of the data does not allow interpretation in terms of relative strengths of activations of certain areas. Parameters such as item duration and frequency strongly influence activation patterns (Price et al., 1994, Price, Moore, and Frackowiak, 1996).

**Table 3.** Brain regions found activated in more than one out of 28 experiments comparing passive listening to a silent control condition. Cells show the number of experiments reporting a given region in relation to the number of experiments covering it (in brackets). Reliable activations are shaded in gray. Following Rademacher et al. (1992), the border between anterior and posterior temporal regions was located at the rostralateral end of the first transverse sulcus, corresponding approximately to  $y = -12$  in the coordinate system of Talairach and Tournoux (1988).

			sentences	words	pseudowords	tones
Frontal lobes	Right	inferior frontal gyrus, pars triangularis	0 (4)	0 (10)	0 (4)	2 (10)
		inferior frontal gyrus, pars opercularis	0 (4)	0 (10)	0 (4)	3 (10)
	Left	inferior frontal gyrus, pars triangularis	0 (4)	2 (10)	0 (4)	0 (10)
		inferior frontal gyrus, pars opercularis	0 (4)	0 (10)	0 (4)	2 (10)
		inferior frontal gyrus, frontoorbital	3 (4)	0 (10)	0 (3)	0 (7)
Temporal lobes	Right	anterior superior temporal gyrus	2 (4)	5 (10)	2 (4)	3 (10)
		posterior superior temporal gyrus	3 (4)	8 (10)	4 (4)	9 (10)
		anterior middle temporal gyrus	3 (4)	5 (10)	0 (4)	0 (10)
		posterior middle temporal gyrus	3 (4)	5 (10)	2 (4)	2 (10)
	Left	anterior superior temporal gyrus	4 (4)	4 (10)	2 (4)	2 (10)
		posterior superior temporal gyrus	4 (4)	8 (10)	4 (4)	10 (10)
		anterior middle temporal gyrus	4 (4)	3 (10)	1 (4)	0 (10)
		posterior middle temporal gyrus	4 (4)	6 (10)	1 (4)	1 (10)

## 2.2. Whole-brain level results

Overall, the brain activations observed during the presentation of auditory stimuli involve primarily the temporal and posterior inferior frontal lobes. Taking into account the complete data set from all 55 experiments, only 13 (5 frontal, 8 temporal) of 108 regions were found to be activated in more than one study. Table 3 summarizes for these 13 regions the findings of the 28 experiments using sentences, words, pseudowords, or tone stimuli compared to silent control conditions. The table lists the number of experiments reporting a certain region as activated in relation to the number of experiments in which this region could have been found given that it was covered by the field-of-view. Regions that by our criterion have been reliably replicated are marked in gray. For auditorily presented sentences and words, these include the bilateral anterior and posterior superior temporal gyri and the bilateral anterior and posterior middle temporal gyri. Sentences and words seem to differ, however, with respect to involvement of the left inferior frontal cortex. Whereas for sentence presentation activation of the left fronto-orbital cortex (BA 47) was reliably replicated, this region was not reported for auditorily presented words. Conversely, the pars triangularis (BA 45) of the left inferior frontal gyrus was found in two of ten experiments on word listening. Based on experiments with silent baseline conditions alone, this proportion is not reliable; however BA 45 was also found in two (Binder et al., 1996; Price et al., 1996) of five experiments using acoustically more complex control conditions. For passive listening to sentences, the pars triangularis was only reported by Schlosser et al. (1998), where the control condition was listening to an unknown language. Frontal regions were not reliably replicated for experiments in which subjects listened to pseudowords or meaningless syllables. Only one study (Binder et al., 2000, comparing pseudowords to tones) found bilateral activation of the pars opercularis (BA 44). Temporal regions reliably found for pseudoword and syllable presentation are the bilateral anterior and posterior superior temporal gyri and the right posterior middle temporal gyrus. In contrast to linguistic stimuli, tone stimuli of varying spectral complexity seem to reliably activate the superior temporal gyri bilaterally, but not the middle temporal gyri. The reports of frontal activations elicited

by tone stimuli agree to a reliable extent on the pars opercularis (BA 44) of the right inferior frontal gyrus.

### **2.3. Fine-grained analysis of the temporal lobes**

In a finer-grained analysis of the patterns of reported activation foci in the temporal lobes, we first present the areas that, to date, can be considered as reliable focal activation areas (Figure 1 (p. 33)) for different kinds of stimuli. This reliability estimate is based on overlap of reported activation maxima from experiments with silent baseline conditions (Table 1, Figure 2 (p. 34)). We then describe the patterns of reported activation maxima obtained in experiments using acoustically or linguistically more complex control conditions (Table 2, Figure 3 (p. 35)).

#### **2.3.1. Sentences**

*Silent control.* Both temporal lobes exhibit an anterior and a posterior region of reliable overlap of activation maxima (red areas in Figure 1). In the  $z$ -dimension (ventral-dorsal), these areas are approximately centered on the superior temporal sulcus. The posterior regions reach dorsally into the primary auditory cortex, but extend for the greater part more ventrally. In the right temporal lobe, the posterior activation maxima cluster (Figure 2, red) less tightly than in the left.

*Acoustic or linguistic control.* The control conditions used in three studies on passive sentence listening (Schlosser et al., 1998: unknown language; Giraud et al., 2000: vowels; Scott et al., 2000: spectrally rotated sentences) shared the acoustic and phonological properties of sentences to different extents, but did not involve semantic or syntactic processing. The resulting patterns of activation foci (Figure 3, red foci) differ in hemisphere-specific ways from those found with a silent baseline (Figure 2, red foci). In the right temporal lobe, no anterior temporal activation maxima are reported, and the maxima tend to cluster around the primary auditory cortex; in the left temporal lobe, roughly the opposite pattern is observed. Two studies (Schlosser et al., 1998; Giraud et al., 2000) report more posterior activation foci along the superior temporal sulcus than found in studies using a silent baseline.

### **2.3.2. Words**

*Silent control* As with sentences, passive listening to words seems to reliably elicit activation foci in two areas of both temporal lobes (Figure 1, yellow). On the left, these two areas are posteriorly adjacent to the sentence areas. On the right, the posterior area largely overlaps with the sentence area and, in addition, includes the primary auditory cortex. Although some word foci (Figure 2, yellow) have been reported as far anterior as sentence foci (Figure 2, red), the area of reliable overlap reaches less rostrally than the sentence area.

*Acoustic or linguistic control.* Activation foci (Figure 3, yellow) shift towards the superior temporal sulcus in both hemispheres. It is mainly on the left that even more ventral activation foci in the middle and inferior temporal gyri are reported (Binder et al., 1996; Price et al., 1996). Similar to the pattern observed for sentences, anterior temporal activations are reported with acoustically complex control conditions on the left (Price et al., 1996) but not on the right.

### **2.3.3. Pseudowords/meaningless syllables**

*Silent control* The reported activation foci overlap reliably in the posterior primary auditory cortex and the adjacent superior temporal gyri in both temporal lobes (Figures 1 and 2, green). These overlap areas are similar to those observed for words. On the left, a second more anterior area of overlap is caudally adjacent to that found for words. More anterior temporal activation of the right superior temporal lobe has been found in one study (Bookheimer et al., 1998), and is, to date, not reliable.

*Acoustic or linguistic control.* The use of control conditions, such as tones (Binder et al., 2000; Jäncke et al., 2002), noise bursts (Jäncke et al., 2002), or vowels (Benson et al., 2001; Jäncke et al., 2002), leads to a shift of activation maxima (Figure 3, green) away from the auditory cortex. Activation foci are more frequently found in the posterior superior temporal planes than in studies of word or sentence listening. In contrast to word listening maxima, activation maxima for pseudowords are only

observed more ventrally than the superior temporal sulcus on the right, not on the left (Jäncke et al., 2002; Benson et al., 2001).

#### **2.3.4. Tones**

*Silent control* Activation maxima cover both auditory cortices and the posteriorly adjacent superior temporal cortices (Figures 1 and 2, blue). This also holds for medial superior temporal lobe activation maxima, which were analyzed separately (not shown in Figure 1).

*Acoustic control.* Comparing frequency- or amplitude-modulated tones to stationary stimuli strongly reduces the number of activation foci (Figure 3, blue) found near the posterior superior temporal planes, suggesting that this part of the superior temporal gyrus does not specifically respond to such modulations. Similar to the linguistic stimuli, activation maxima for non-linguistic auditory stimuli tend to be observed nearer to the superior temporal sulcus when compared to tone stimuli than when compared to silence. However, their reported location tends to be less ventral than that of linguistic stimuli, and to respect the superior temporal sulcus as a ventral border in both hemispheres.

### **3. Structure-function relationships**

#### **3.1. General observations**

Our synopsis suggests that no brain areas other than the posterior inferior frontal and the temporal cortex are reliably activated during passive listening to speech. As has been previously observed (Mazoyer et al., 1993; Binder et al., 1996), the temporal cortex activation *areas* for different kinds of auditory stimuli seem to show roughly the subset-superset relationships one might expect from the hierarchical organisation of the acoustic, phonological, lexical semantic, and syntactic stimulus properties (see Table 3). All auditory stimuli reliably activate the central and posterior parts of the superior temporal gyri. With increasing linguistic complexity, more ventral and anterior parts of the temporal lobes are recruited. This hierarchical relationship is not

as clearly observed in the frontal lobes. For example, the right posterior inferior frontal gyrus, which is reliably activated by tone stimuli, is typically not activated by pseudoword, word, and sentence stimuli, although the latter kinds of stimuli have, besides their additional linguistic properties, many properties in common with the simpler acoustic stimuli.

Our meta-analysis of the overlap of activation *maxima* (as opposed to total activation areas) shows that there is also a systematic deviation from this expected subset-superset relationship in the temporal lobes. Rather than showing ever larger areas of reliable focal activation as one proceeds from non-linguistic stimuli to stimuli with increasingly linguistic properties, the hemodynamic activation maxima observed for different kinds of auditory stimuli are clustered in a stimulus-dependent way on the lateral temporal plane of both hemispheres. This is most clearly seen in the anterior left temporal lobe, where the activation maxima for sentences, words, pseudowords, and tones are ordered along an anterior-to-central dimension. In the posterior left temporal lobe, activation maxima for pseudowords and syllables but not for words are reported in the superior temporal plane. Activation maxima for words but not for sentences are reported in the posterior middle/inferior temporal gyri. Considering that words share all the acoustic and linguistic properties of pseudowords, and sentences all the acoustic and linguistic properties of words, these patterns of reported activation maxima do not seem to follow a “subtraction logic” according to which foci observed for listening to sentences compared to a silent baseline should reflect acoustic or phonological processing as well as semantic or syntactic processing. These findings suggest that the brain response to a particular auditory stimulus property is not uniform, but actively adapted to the other properties of the stimulus. They can only be accounted for by assuming that those neuronal populations of the temporal cortex that are sensitive to the ‘highest’ linguistic property of an auditory stimulus show the strongest and most consistent hemodynamic response, thus determining the location of the statistical activation maxima. According to this reasoning, the stimulus-specific patterns of reported activation maxima in passive listening experiments with a silent baseline condition provide information on the stimulus properties that particular cortical areas are sensitive to. In

the following sections, we will attempt to draw some tentative conclusions about the temporal cortex regions subserving different levels of linguistic processing. To this end we will also take into account the spatial distribution of activation foci found in comparisons with control conditions other than silence, as well as electrophysiological, MEG, and clinical data.

### **3.2. Acoustic processing**

In both hemispheres, activation maxima for non-linguistic acoustic stimuli compared to silence are reported in the primary auditory cortex and the posterior superior temporal plane, suggesting a role of these regions in acoustic processing. Activation of these areas starts already 20-80 ms after the onset of pure tone stimuli (Yvert et al., 2001). The precise location of activation maxima within these regions depends on acoustic properties such as sound pressure (Bilecen et al., 2002) and frequency (Lauter et al., 1985; Wessinger et al., 1997; Bilecen et al., 1998; Talavage et al., 2000; Le et al., 2001; Engelien et al., 2002; Schönwiesner, von Cramon, and RübSamen, 2002). Activation of these regions is rarely observed for words or sentences when compared to tones, so that these regions do not seem to be related to word level processing (Binder et al., 1996). When spectrally or temporally more complex acoustic stimuli are compared to simpler ones, more ventral activation foci near the central (on the left also posterior) superior temporal sulcus are observed, so that a role of this part of the superior temporal sulci in acoustic processing is probable.

### **3.3. Phonetic/phonological processing**

At present, neuroimaging data do not allow for a distinction between phonetic and phonological processing. Throughout this paper, we use the term 'phonological processing' in the loose sense of a processing level operating on discrete categories. Although it is as yet unclear whether and how lexical access indeed involves discrete intermediate representations (see section 1), there is clear evidence for categorical perception of speech sounds and for language-specific phonological constraints in speech processing (see Phillips, 2001, for a discussion of acoustic, phonetic and phonological processing levels).

Pseudowords have all the properties of real words except for meaning. One might therefore assume phonological processing to be involved in activation observed with pseudoword presentation. At the same time, though, pseudowords are also acoustically more complex than most control stimuli that have been used (see Scott and Wise, 2003, for a detailed discussion of different control stimuli), so that stronger activation of a neural population for pseudowords as compared to non-linguistic stimuli is a necessary but not a sufficient condition for a role of this population in phonological processing. For this reason, we will have to consider additional evidence to achieve a tentative conclusion as to the neural structures involved in phonological processes.

In both hemispheres, activation maxima for pseudowords are reported in the central to posterior dorsal aspect of the superior temporal gyri (Wernicke's area), when compared to silent rest. Activation of the bilateral posterior superior temporal plane and the right central middle temporal gyrus by monosyllabic pseudowords does not seem to be cancelled out by tone, noise or phonologically simpler control stimuli (Benson et al., 2001; Jäncke et al., 2002), suggesting a function in the processing of complex acoustic properties underlying linguistic stimuli or a phonological processing function. Furthermore, it is possible that the right and left temporal lobes differ in the kind of processing they support. Left hemispheric dominance for phonemes as opposed to non-linguistic acoustic stimuli has been observed in a number of mismatch negativity (MMN), mismatch field (MMF) and dipole modelling studies (e.g. Näätänen et al., 1997; Alho et al., 1998; Gootjes et al., 1999; Rinne et al., 1999; Szymanski, Rowley and Roberts, 1999; Szymanski et al., 2001; also see Shtyrov et al., 1999, for converging evidence). While these findings suggest some specialisation for language stimuli in the left temporal lobe, they do not prove phonological processing in the left superior temporal lobe nor exclude phonological processing in the right temporal lobe (see Phillips, 2001, for detailed discussion of electrophysiological findings).

To distinguish between acoustic and phonological processing, Phillips et al. (2000) exploited phonemic category perception effects in an elegantly designed MEG study. Using stimuli from a /tæ-dæ/ voice onset time (VOT) continuum, they elicited a

MMF response to infrequent /t/ phonemes with long VOTs when frequent stimuli with shorter VOTs varied acoustically within the phonemic category /d/ but not when the same amount of acoustic variation occurred across the phonemic /d-t/ boundary, preventing the perception of a standard phoneme. The MMF response occurred in a time window of 150-200 ms post stimulus onset. The response had its source in the left auditory cortex (the right temporal lobe was not measured). In a near-infrared spectroscopy (NIRS) experiment, Minagawa-Kawai et al. (2002) demonstrated hemodynamic activation of the left auditory cortex and the adjacent planum temporale for between-category compared to equidistant within-category vowel-length contrasts. The effect was bilateral in some subjects, but never right-lateralized. Although it is debatable whether the observed category effects reflect phonological rather than phonetic categories, the effects clearly show that the processing of linguistic stimuli in the left auditory cortex, possibly also the adjacent posterior temporal plane and right homologue areas, goes beyond acoustic representations. Further electrophysiological evidence for phonological processing comes from a study by Dehaene-Lambertz, Dupoux, and Gout (2000), who showed that MMN responses can be sensitive to language-specific phonotactic constraints.

Note that the set of phonetic features studied so far is limited and does not allow the conclusion that all phonetically relevant distinctions are processed in the temporal lobes. In a recent fMRI study, Mathiak et al. (2002) compared the hemodynamic responses to two categorical perception tasks in which the identity of a word-medial consonant was either signalled by voice onset time or by the length of the preceding word-medial pause (closure time, CLT). In direct comparisons, they found stronger activation of the left supratemporal plane in the VOT condition, and stronger frontal and cerebellar activation in the CLT condition. The latter finding confirmed clinical findings in patients with cerebellar atrophy, who did not show categorical perception for the CLT contrast (Ackermann et al., 1997). Clinical observations also suggest an involvement of the posterior superior temporal lobes in phonological processing. Bilateral lesions of the posterior superior temporal lobes (in some cases also left unilateral lesions, see Griffiths, Rees, and Green, 1999) may result in word deafness, an impairment of language perception with relatively spared

processing of non-linguistic auditory stimuli (Buchman et al., 1986; Griffiths, Rees, and Green, 1999; Poeppel, 2001). Unilateral lesions of the dominant posterior superior temporal gyrus, as in Wernicke's aphasia, may cause more subtle phonological processing deficits, such as problems with the discrimination of phonological contrasts, in particular place of articulation (Blumstein, 1995).

In sum, there is good evidence for a role of the central to posterior superior temporal gyri in phonetic or phonological processing. The left posterior superior temporal gyrus seems to be dominant. In contrast to the robust findings from studies using phoneme discrimination or monitoring tasks (Demonet et al., 1992, 1996) there is, to date, little evidence for an involvement of the left posterior frontal cortex in phonological processing during passive listening.

### **3.4. Lexical processing**

Words and pseudowords activate wholly or partially matching candidate words; non-linguistic stimuli do not. Activation of a word form can also make available conceptual information, so that syntactic and semantic properties of multiple lexical candidates may also be simultaneously active. Brain areas that are activated by pseudowords may thus to some extent reflect conceptual processing. Nonetheless, words differ from pseudowords in that there is a winner of the lexical competition process, and the successful retrieval of a lexical entry is necessary for word recognition to occur. It may therefore be assumed that brain areas that are activated more strongly for words than for pseudowords reflect additional syntactic and semantic processes involved in or subsequent to lexical recognition.

Activation foci for all linguistic stimuli but not for non-linguistic stimuli are mainly reported for anterior and posterior parts of the left superior temporal sulcus (see also Giraud and Price, 2001), both compared to silent rest and to acoustic control conditions. These areas may therefore be considered as candidates for involvement in lexical access and competition processes. There is evidence that lexical access may be affected in patients with fluent aphasia, many of whom have lesions of the posterior part of the left superior temporal sulcus. In such patients, Milberg, Blumstein, and Dworetzky (1988) observed enhanced semantic facilitation by phonologically

distorted primes (not only ‘cat’, but also the pseudoword ‘wat’ primed ‘dog’). Since these patients can indeed discriminate ‘cat’ from ‘wat’, Blumstein and Milberg (2000) locate the effect at the lexical level, assuming a lower threshold for the activation of lexical entries or a general “overactivation” of lexical entries.

Regions in which activation maxima are found for words but not for pseudowords are the most likely candidates for a role in the retrieval of lexically stored information after resolution of the competition process. Compared to silent rest, reliable word activation maxima are found more ventrally and anteriorly than pseudoword maxima in both hemispheres. Due to a broader distribution of pseudoword foci, these location differences disappear with acoustic control conditions in the right hemisphere, but not in the left. Here word activation foci are found more anteriorly (Price et al., 1996; Mummery et al., 1999) and postero-ventrally (Binder et al., 1996; Price et al., 1996) than pseudoword foci. A possible relation of these areas to semantic processing has been suggested by a number of authors (for an overview see Price, Indefrey, and van Turenhout, 1999). More recently, Scott et al. (2000) assigned a particular role for the processing of “intelligible” speech to a left anterior superior temporal pathway, whereas Hickok and Poeppel (2000) suggested that the auditory-conceptual interface involves a posterior pathway to the left temporal-parietal-occipital junction. Rather than supporting an exclusive role of anterior or posterior temporal structures in semantic processing, the data reviewed here suggest that different parts of the left middle and inferior temporal gyri may be involved in lexical semantic (but possibly also lexical syntactic) processing during passive listening. In addition, Broca's area, in particular BA 45, seems to be involved in the retrieval of lexical information (Binder et al., 1996; Price et al., 1996). This rather broad distribution of brain areas associated with lexical processing is also reflected in the heterogeneity of clinical findings. Auditory language comprehension deficits in Wernicke's aphasia do not seem to be strongly related to phonological deficits (Blumstein, 1995; Hickok and Poeppel, 2000). In patients with impaired comprehension, the lesions are typically not confined to the left superior temporal gyrus but extend ventrally and caudally into the middle temporal gyrus and the inferior parietal lobe (Damasio, 1992). Impairments of language comprehension with

a relatively preserved ability to repeat heard words, suggesting intact word form representations, are found in transcortical sensory aphasia (TCSA) and semantic dementia. TCSA may result from lesions of the left temporo-parietal or anterior temporal/inferior frontal cortex (Damasio, 1991; Berthier, 1999). Semantic dementia is a neurodegenerative disease affecting the left anterior and inferior medial temporal cortex. While the observed impairments suggest that the left anterior temporal lobe is necessary for language comprehension, Scott and Wise (2003) point out that surgical removal of this region in epileptic patients does not seem to result in a semantic processing deficit (see also Hagoort et al., 1999). Although the neural substrate of language functions may of course have been altered in epileptic patients, the observation suggests that the TCSA and dementia data should be interpreted with caution.

In sum, the available evidence suggests that lexical processing involves anterior and posterior parts of the left superior temporal sulcus. Based on the differences observed between word and pseudoword activation patterns it can be assumed that there are additional lexical retrieval processes for items that are recognized (win the lexical competition). These processes seem to recruit the left posterior inferior frontal gyrus as well as more anterior and postero-ventral temporal areas reaching into the middle and inferior temporal gyri. It should be noted that this picture is almost certainly incomplete considering the available data on category-specific semantic representations which were not targeted by the experiments analyzed here (see Caramazza and Shapiro, this volume).

### **3.5. Sentence level processing**

Reliable focal activation areas for sentences but not words compared to silent rest are the temporal poles and the central regions of the middle temporal gyri as well as the left posterior inferior frontal gyrus. With acoustic or phonological control conditions, left temporal pole activation has been confirmed by Scott et al. (2000), whereas the other regions are either no longer found or found for other stimuli as well. These findings suggest the left temporal pole as the best candidate area for sentence-level processing in the temporal lobes. They do not allow a decision as to whether syntax,

sentence-level meaning, or prosody is processed. Mazoyer et al. (1993), who were the first to observe bilateral temporal pole activations for sentences, found similar activations for meaningless pseudoword and syntactic prose sentences. Friederici, Meyer, and von Cramon (2000) also report bilateral anterior temporal lobe activation, although more dorsally, for normal and pseudoword sentences. These findings suggest a syntactic rather than semantic function - a view that is also advocated by Dronkers (1994) on the basis of clinical data. For aphasic patients with syntactic processing impairments (agrammatism) she found a common lesion area in the anterior temporal lobe that was spared in aphasic patients, who were not agrammatic. By contrast, Hagoort et al. (1999) did not find any syntactic deficit in epileptic patients who underwent anterior temporal lobe surgery. Such conflicting findings again point to the necessity of interpreting structure-function relations in clinical populations with caution given that neural reorganisation may have occurred to an unknown extent. If the left anterior temporal pole indeed has a syntactic function, one would expect this area to be more strongly activated for sentences than for words. Wong et al. (1999), however, did not find temporal pole activation but posterior inferior frontal (BA 47) activation in a direct comparison of sentences and words. The latter area was also found to be reliably activated in our comparison of sentences to silent rest, so that it probably subserves a syntactic or semantic sentence-level processing function. BA 47 is ventrally adjacent to Broca's area (BA 44/45), which is typically found to be active in experiments using syntactic violation or judgement tasks (see Kaan and Swaab, 2002, as well as Friederici, this volume, for an overview of frontal and temporal activations observed with such paradigms). Considering that persisting syntactic processing deficits seem to require frontal lesions that go beyond BA 44 and 45 (Mohr et al., 1978), BA 47 may well play a role in syntactic processing. Considering further that syntactic processing deficits can also occur after posterior superior temporal lesions (Caplan, Hildebrand, and Makris, 1996), neither the left anterior temporal lobe nor the left inferior frontal lobe seem sufficient for syntactic natural language processing.

In sum, there is evidence suggesting involvement of the left temporal pole and the left posterior inferior frontal gyrus in sentence-level processing. To date, however, the exact role of these two areas is unclear.

#### **4. Conclusions**

Auditory language input is processed in a left posterior frontal and bilateral temporal cortical network. Within this network, no processing level is related to a single cortical area. Although auditory language processing activates both temporal lobes, they are not equipotential. In the right temporal lobe, activation foci for different auditory stimuli are found in largely overlapping areas and there is to date no clear evidence for hemodynamic activation related to postlexical linguistic processing. In the left temporal lobe, the reported activation foci for different kinds of auditory stimuli show clearly distinguishable patterns, and areas that seem to be specialized for phonological (posterior superior temporal gyrus), lexical (anterior and posterior superior temporal sulcus/middle temporal gyrus, posterior inferior frontal gyrus), and sentence-level (temporal pole, posterior inferior frontal gyrus) processing can be identified with some confidence on the basis of the available evidence.

#### **Acknowledgements**

We thank Frauke Hellwig for making the illustrations, and Peter Hagoort and James McQueen for helpful comments on the text.

#### **References**

- Ackermann, H., S. Graber, I. Hertrich, and I. Daum, 1997. Categorical speech perception in cerebellar disorders. *Brain Lang.* 60:323-331.
- Alho, K., J. F. Connolly, M. Cheour, A. Lehtokoski, M. Huotilainen, J. Virtanen, R. Aulanko, and R. J. Ilmoniemi, 1998. Hemispheric lateralization in preattentive processing of speech sounds. *Neurosci. Lett.* 258:9-12.

- Belin, P., R. J. Zatorre, R. Hoge, A. C. Evans, and B. Pike, 1999. Event-related fMRI of the auditory cortex. *Neuroimage* 10:417-429.\*
- Belin, P., M. Zilbovicius, S. Crozier, L. Thivard, and A. Fontaine, 1998. Lateralization of speech and auditory temporal processing. *J. Cogn. Neurosci.* 10:536-540.\*
- Benson, R. R., D. H. Whalen, M. Richardson, B. Swainson, V. P. Clark, S. Lai, and A. M. Liberman, 2001. Parametrically dissociating speech and nonspeech perception in the brain using fMRI. *Brain Lang.* 78:364-396.\*
- Berthier, M. *Transcortical aphasias*. Hove: Psychology Press, 1999.
- Bilecen, D., K. Scheffler, N. Schmid, K. Tschopp, and J. Seelig, 1998. Tonotopic organization of the human auditory cortex as detected by BOLD-FMRI. *Hear. Res.* 126:19-27.
- Bilecen, D., E. Seifritz, K. Scheffler, J. Henning, and A. C. Schulte, 2002. Amplitude of the human auditory cortex: An fMRI study. *Neuroimage* 17:710-718.
- Binder, J. R., J. A. Frost, T. A. Hammeke, P. S. F. Bellgowan, S. M. Rao, and R. W. Cox, 1999. Conceptual processing during the conscious resting state: A functional MRI study. *J. Cogn. Neurosci.* 11:80-93.
- Binder, J. R., J. A. Frost, T. A. Hammeke, P. S. F. Bellgowan, J. A. Springer, J. N. Kaufman, and E. T. Possing, 2000. Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10:512-528.\*
- Binder, J. R., J. A. Frost, T. A. Hammeke, S. M. Rao, and R. W. Cox, 1996. Function of the left planum temporale in auditory and linguistic processing. *Brain* 119:1239-1247.\*
- Blumstein, S. E., 1995. The neurobiology of the sound structure of language. In M. S. Gazzaniga, ed. *The cognitive neurosciences*. Cambridge, MA: MIT Press, pp. 915-930.
- Blumstein, S. E., and W. P. Milberg, 2000. Language deficits in Broca's and Wernicke's aphasia: A singular impairment. In Y. Grodzinsky, L. P. Shapiro, and D. Swinney, eds. *Language and the Brain*. San Diego, CA: Academic Press, pp. 167-183.

- Bookheimer, S. Y., T. A. Zeffiro, T. A. Blaxton, W. D. Gaillard, B. Malow, and W. H. Theodore, 1998. Regional cerebral blood flow during auditory responsive naming: evidence for cross-modality neural activation. *Neuroreport* 9:2409-2413.\*
- Buchman, A. S., D. C. Garron, J. E. Trostcardamone, M. D. Wichter, and M. Schwartz, 1986. Word deafness - 100 years later. *J. Neurol. Neurosurg. Psychiatry* 49:489-499.
- Caplan, D., N. Alpert, and G. Waters, 1998. Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *J. Cogn. Neurosci.* 10:541-552.
- Caplan, D., N. Alpert, G. Waters, and A. Olivieri, 2000. Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Hum. Brain Mapp.* 9:65-71.
- Caplan, D., N. Hildebrandt, and N. Makris, 1996. Location of lesions in stroke patients with deficits in syntactic processing in sentence comprehension. *Brain* 119:933-949.
- Celsis, P., K. Boulanouar, B. Doyon, J. P. Ranjeva, I. Berry, J. L. Nespoulous, and F. Chollet, 1999. Differential fMRI responses in the left posterior superior temporal gyrus and left supramarginal gyrus to habituation and change detection in syllables and tones. *Neuroimage* 9:135-144.\*
- Connine, C. M., D. G. Blasko, and J. Wang, 1994. Vertical similarity in spoken word recognition: multiple lexical activation, individual differences, and the role of sentence context. *Percept. Psychophys.* 56:624-636.
- Damasio, A. R., 1992. Aphasia. *N. Engl. J. Med.* 326:531-539.
- Damasio, H., 1991. Neuroanatomical correlates of the aphasias. In M. Taylor Sarno, ed. *Acquired aphasia*. New York: Academic Press, pp. 45-71.
- Dehaene-Lambertz, G., E. Dupoux, and A. Gout, 2000. Electrophysiological correlates of phonological processing: A cross-linguistic study. *J. Cogn. Neurosci.* 12:635-647.

- Demonet, J. F., F. Chollet, S. Ramsay, D. Cardebat, J. L. Nespoulous, R. Wise, A. Rascol, and R. Frackowiak, 1992. The anatomy of phonological and semantic processing in normal subjects. *Brain* 115:1753-68.
- Demonet, J. F., J. A. Fiez, E. Paulesu, S. E. Petersen, and R. J. Zatorre, 1996. Pet studies of phonological processing - a critical reply to Poeppel. *Brain Lang.* 55:352-379.
- Di Salle, F., E. Formisano, E. Seifritz, D. E. J. Linden, K. Scheffler, C. Saulino, G. Tedeschi, F. E. Zanella, A. Pepino, R. Goebel, and E. Marciano, 2001. Functional fields in human auditory cortex revealed by time-resolved fMRI without interference of EPI noise. *Neuroimage* 13:328-338.\*
- Dronkers, N. F., D. P. Wilkins, R. D. Vanvalin, B. B. Redfern, and J. J. Jaeger, 1994. A reconsideration of the brain-areas involved in the disruption of morphosyntactic comprehension. *Brain Lang.* 47:461-463.
- Engelien, A., Y. H. Yang, W. Engelien, J. Zonana, E. Stern, and D. A. Silbersweig, 2002. Physiological mapping of human auditory cortices with a silent event-related fMRI technique. *Neuroimage* 16:944-953.\*
- Fiez, J. A., M. E. Raichle, D. A. Balota, P. Tallal, and S. E. Petersen, 1996. Pet activation of posterior temporal regions during auditory word presentation and verb generation. *Cereb. Cortex* 6:1-10.\*
- Friederici, A. D., M. Meyer, and D. Y. von Cramon, 2000. Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang.* 74:289-300.
- Giraud, A. L., and C. J. Price, 2001. The constraints functional neuroimaging places on classical models of auditory word processing. *J. Cogn. Neurosci.* 13:754-765.
- Giraud, A. L., E. Truy, R. S. J. Frackowiak, M. C. Gregoire, J. F. Pujol, and L. Collet, 2000. Differential recruitment of the speech processing system in healthy subjects and rehabilitated cochlear implant patients. *Brain* 123:1391-1402.\*
- Gootjes, L., T. Raiji, R. Salmelin, and R. Hari, 1999. Left-hemisphere dominance for processing of vowels: a whole-scalp neuromagnetic study. *Neuroreport* 10:2987-2991.

- Gow, D. W., and P. C. Gordon, 1995. Lexical and prelexical influences on word segmentation - evidence from priming. *J. Exp. Psychol.-Hum. Percept. Perform.* 21:344-359.
- Griffiths, T. D., A. Rees, and G. G. R. Green, 1999. Disorders of human complex sound processing. *Neurocase* 5:365-378.
- Grosjean, F., and U. H. Frauenfelder, eds. *A guide to spoken word recognition paradigms*. Hove: Psychology Press, 1997. (also a special issue of *Lang. Cogn. Process.* 6, Volume 11, 1996).
- Hagoort, P., N. Ramsey, G. J. Rutten, and P. van Rijen, 1999. The role of the left anterior temporal cortex in language processing. *Brain Lang.* 69:322-325.
- Hall, D. A., I. S. Johnsrude, M. P. Haggard, A. R. Palmer, M. A. Akeroyd, and A. Q. Summerfield, 2002. Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 12:140-149.\*
- Hickok, G., and D. Poeppel, 2000. Towards a functional neuroanatomy of speech perception. *Trends Cogn. Sci.* 4:131-138.
- Holcomb, H. H., D. R. Medoff, P. J. Caudill, Z. Zhao, A. C. Lahti, R. F. Dannals, and C. A. Tamminga, 1998. Cerebral blood flow relationships associated with a difficult tone recognition task in trained normal volunteers. *Cereb. Cortex* 8:534-542.\*
- Indefrey, P., F. M. Hellwig, H. Herzog, R. Seitz, and P. Hagoort, 2003. Neural responses to the production and comprehension of syntax in identical utterances. *Brain Lang.* (to appear).
- Indefrey, P., and W. J. M. Levelt, 2000. The neural correlates of language production. In M. S. Gazzaniga, ed. *The new cognitive neurosciences*. Cambridge, MA: MIT Press, pp. 845-865.
- Indefrey, P., and W. J. M. Levelt, 2003. The spatial and temporal signatures of word production components. *Cognition* (to appear).
- Jäncke, L., S. Mirzazade, and N. J. Shah, 1999. Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. *Neurosci. Lett.* 266:125-128.\*

- Jäncke, L., T. Wüstenberg, H. Scheich, and H. J. Heinze, 2002. Phonetic perception and the temporal cortex. *Neuroimage* 15:733-746.\*
- Kaan, E., and T. Y. Swaab, 2002. The brain circuitry of syntactic comprehension. *Trends Cogn. Sci.* 6:350-356.
- Lauter, J. L., P. Herscovitch, C. Formby, and M. E. Raichle, 1985. Tonotopic organization in human auditory cortex revealed by Positron Emission Tomography. *Hear. Res.* 20:199-205.
- Le, T. H., S. Patel, and T. P. L. Roberts, 2001. Functional MRI of human auditory cortex using block and event-related designs. *Magn. Reson. Med.* 45:254-260.
- Lockwood, A. H., R. J. Salvi, M. L. Coad, S. A. Arnold, D. S. Wack, B. W. Murphy, and R. F. Burkard, 1999. The functional anatomy of the normal human auditory system: Responses to 0.5 and 4.0 kHz tones at varied intensities. *Cereb. Cortex* 9:65-76.\*
- Luce, P. A., and E. A. Lyons, 1999. Processing lexically embedded spoken words. *J. Exp. Psychol.-Hum. Percept. Perform.* 25:174-183.
- Marslen-Wilson, W., and P. Warren, 1994. Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychol. Rev.* 101:653-675.
- Mathiak, K., I. Hertrich, W. Grodd, and H. Ackermann, 2002. Cerebellum and speech perception: A functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 14:902-912.
- Mazoyer, B. M., N. Tzourio, V. Frak, A. Syrota, N. Murayama, O. Levrier, G. Salamon, S. Dehaene, L. Cohen, and J. Mehler, 1993. The cortical representation of speech. *J. Cogn. Neurosci.* 5:467-479.
- McGuire, P. K., D. A. Silbersweig, and C. D. Frith, 1996. Functional neuroanatomy of verbal self-monitoring. *Brain* 119:907-917.
- McQueen, J. M., D. Dahan, and A. Cutler, 2003. Continuity and gradedness in speech processing. In N. O. Schiller, and A. S. Meyer, eds. *Phonetics and Phonology in Language Comprehension and Production. Differences and Similarities*. Berlin, New York: Mouton de Gruyter, pp. 37-76.

- McQueen, J. M., D. Norris, and A. Cutler, 1994. Competition in spoken word recognition - spotting words in other words. *J. Exp. Psychol.-Learn. Mem. Cogn.* 20:621-638.
- McQueen, J. M., D. Norris, and A. Cutler, 1999. Lexical influence in phonetic decision making: Evidence from subcategorical mismatches. *J. Exp. Psychol.-Hum. Percept. Perform.* 25:1363-1389.
- Mellet, E., N. Tzourio, F. Crivello, M. Joliot, M. Denis, and B. Mazoyer, 1996. Functional anatomy of spatial mental imagery generated from verbal instructions. *J. Neurosci.* 16:6504-6512.\*
- Milberg, W., S. Blumstein, and B. Dworetzky, 1988. Phonological processing and lexical access in aphasia. *Brain Lang.* 34:279-293.
- Minagawa-Kawai, Y., K. Mori, I. Furuya, R. Hayashi, and Y. Sato, 2002. Assessing cerebral representations of short and long vowel categories by NIRS. *Neuroreport* 13:581-584.
- Mirz, F., T. Ovesen, K. Ishizu, P. Johannsen, S. Madsen, A. Gjedde, and C. B. Pedersen, 1999. Stimulus-dependent central processing of auditory stimuli - A PET study. *Scand. Audiol.* 28:161-169.\*
- Mohr, J. P., M. S. Pessin, S. Finkelstein, H. H. Funkenstein, G. W. Duncan, and K. R. Davis, 1978. Broca aphasia - pathologic and clinical. *Neurology* 28:311-324.
- Müller, B. W., M. Jüptner, W. Jentzen, and S. P. Müller, 2002. Cortical activation to auditory mismatch elicited by frequency deviant and complex novel sounds: A PET study. *Neuroimage* 17:231-239.\*
- Müller, R. A., R. D. Rothemel, M. E. Behen, O. Muzik, T. J. Mangner, and H. T. Chugani, 1997. Receptive and expressive language activations for sentences: A PET study. *Neuroreport* 8:3767-3770.\*
- Mummary, C. J., J. Ashburner, S. K. Scott, and R. J. S. Wise, 1999. Functional neuroimaging of speech perception in six normal and two aphasic subjects. *J. Acoust. Soc. Am.* 106:449-457.\*
- Näätänen, R., A. Lehtokoski, M. Lennes, M. Cheour, M. Huotilainen, A. Iivonen, M. Vainio, P. Alku, R. J. Ilmoniemi, A. Luuk, J. Allik, J. Sinkkonen, and K. Alho,

1997. Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385:432-434.
- Norris, D., J. M. McQueen, and A. Cutler, 1995. Competition and segmentation in spoken word recognition. *J. Exp. Psychol.-Learn. Mem. Cogn.* 21:1209-1228.
- Norris, D., J. M. McQueen, and A. Cutler, 2000. Merging information in speech recognition: Feedback is never necessary. *Behav. Brain Sci.* 23:299-325.
- Norris, D., and R. Wise, 2000. The study of prelexical and lexical processes in comprehension: Psycholinguistics and functional neuroimaging. In M. S. Gazzaniga, ed. *The new cognitive neurosciences*. Cambridge, MA: MIT Press, pp. 867-880.
- Penhune, V. B., R. J. Zatorre, J. D. Macdonald, and A. C. Evans, 1996. Interhemispheric anatomical differences in human primary auditory cortex - probabilistic mapping and volume measurement from magnetic resonance scans. *Cereb. Cortex* 6:661-672.
- Petersen, S. E., P. T. Fox, M. I. Posner, M. Mintun, and M. E. Raichle, 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331:585-9.\*
- Phillips, C., 2001. Levels of representation in the electrophysiology of speech perception. *Cogn. Sci.* 25:711-731.
- Phillips, C., T. Pellathy, A. Marantz, E. Yellin, K. Wexler, D. Poeppel, M. McGinnis, and T. Roberts, 2000. Auditory cortex accesses phonological categories: An MEG mismatch study. *J. Cogn. Neurosci.* 12:1038-1055.
- Poeppel, D., 2001. Pure word deafness and the bilateral processing of the speech code. *Cogn. Sci.* 25:679-693.
- Price, C., P. Indefrey, and M. van Turennout, 1999. The neural architecture underlying the processing of written and spoken word forms. In C. M. Brown, and P. Hagoort, eds. *The neurocognition of language*. Oxford: Oxford University Press, pp. 211-240.
- Price, C. J., C. J. Moore, and R. S. J. Frackowiak, 1996. The effect of varying stimulus rate and duration on brain activity during reading. *Neuroimage* 3:40-52.

- Price, C. J., R. J. S. Wise, E. A. Warburton, C. J. Moore, D. Howard, K. Patterson, R. S. J. Frackowiak, and K. J. Friston, 1996. Hearing and saying - the functional neuro-anatomy of auditory word processing. *Brain* 119:919-931.\*
- Price, C. J., R. J. S. Wise, J. D. G. Watson, K. Patterson, D. Howard, and R. S. J. Frackowiak, 1994. Brain activity during reading - the effects of exposure duration and task. *Brain* 117:1255-1269.
- Rademacher, J., A. M. Galaburda, D. N. Kennedy, P. A. Filipek, and V. S. Caviness, 1992. Human cerebral cortex - localization, parcellation, and morphometry with magnetic-resonance-imaging. *J. Cogn. Neurosci.* 4:352-374.
- Rinne, T., K. Alho, P. Alku, M. Holi, J. Sinkkonen, J. Virtanen, O. Bertrand, and R. Näätänen, 1999. Analysis of speech sounds is left-hemisphere predominant at 100-150 ms after sound onset. *Neuroreport* 10:1113-1117.
- Schlosser, M. J., N. Aoyagi, R. K. Fulbright, J. C. Gore, and G. McCarthy, 1998. Functional MRI studies of auditory comprehension. *Hum. Brain Mapp.* 6:1-13.\*
- Schönwiesner, M., D. Y. von Cramon, and R. Rübsamen, 2002. Is it tonotopy after all? *Neuroimage* 17:1144-1161.
- Scott, S., and R. Wise, 2003. The functional neuroanatomy of prelexical processing in speech perception. : to appear.
- Scott, S. K., C. C. Blank, S. Rosen, and R. J. Wise, 2000. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 12:2400-6.\*
- Shtyrov, Y., T. Kujala, R. J. Ilmoniemi, and R. Näätänen, 1999. Noise affects speech-signal processing differently in the cerebral hemispheres. *Neuroreport* 10:2189-2192.
- Shulman, G., M. Corbetta, R. Buckner, J. Fiez, F. Miezin, M. Raichle, and S. Petersen, 1997. Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J. Cogn. Neurosci.* 9:648-663.
- Soto-Faraco, S., N. Sebastián-Gallés, and A. Cutler, 2001. Segmental and suprasegmental mismatch in lexical access. *J. Mem. Lang.* 45:412-432.
- Streeter, L. A., and G. N. Nigro, 1979. The role of medial consonant transitions in word perception. *J. Acoust. Soc. Am.* 65:1533-1541.

- Stromswold, K., D. Caplan, N. Alpert, and S. Rauch, 1996. Localization of syntactic comprehension by positron emission tomography. *Brain Lang.* 52:452-473.
- Suzuki, M., H. Kitano, T. Kitanishi, R. Ito, A. Shiino, Y. Nishida, Y. Yazawa, F. Ogawa, and K. Kitajima, 2002a. Cortical and subcortical activation with monaural monosyllabic stimulation by functional MRI. *Hear. Res.* 163:37-45.\*
- Suzuki, M., H. Kouzaki, R. Ito, A. Shiino, Y. Nishida, and H. Kitano, 2002b. Cortical activation patterns in functional magnetic resonance images following monaural pure tone stimulation. *Neurosci. Res. Commun.* 30:197-206.\*
- Szymanski, M. D., D. W. Perry, N. M. Gage, H. A. Rowley, J. Walker, M. S. Berger, and T. P. L. Roberts, 2001. Magnetic source imaging of late evoked field responses to vowels: toward an assessment of hemispheric dominance for language. *J. Neurosurg.* 94:445-453.
- Szymanski, M. D., H. A. Rowley, and T. P. L. Roberts, 1999. A hemispherically asymmetrical MEG response to vowels. *Neuroreport* 10:2481-2486.
- Tabossi, P., C. Burani, and D. Scott, 1995. Word identification in fluent speech. *J. Mem. Lang.* 34:440-467.
- Talairach, J., and P. Tournoux. *Co-planar stereotaxic atlas of the human brain.3-dimensional proportional system : an approach to cerebral imaging*. Stuttgart: Thieme, 1988.
- Talavage, T. M., P. J. Ledden, R. R. Benson, B. R. Rosen, and J. R. Melcher, 2000. Frequency-dependent responses exhibited by multiple regions in human auditory cortex. *Hear. Res.* 150:225-244.
- Thivard, L., P. Belin, M. Zilbovicius, J. B. Poline, and Y. Samson, 2000. A cortical region sensitive to auditory spectral motion. *Neuroreport* 11:2969-2972.\*
- Vroomen, J., and B. de Gelder, 1997. Activation of embedded words in spoken word recognition. *J. Exp. Psychol.-Hum. Percept. Perform.* 23:710-720.
- Vroomen, J., and B. de Gelder, 1995. Metrical segmentation and lexical inhibition in spoken word recognition. *J. Exp. Psychol.-Hum. Percept. Perform.* 21:98-108.

- Warburton, E., R. J. S. Wise, C. J. Price, C. Weiller, U. Hadar, S. Ramsay, and R. S. J. Frackowiak, 1996. Noun and verb retrieval by normal subjects studies with PET. *Brain* 119:159-179.\*
- Wessinger, C. M., M. H. Buonocore, C. L. Kussmaul, and G. R. Mangun, 1997. Tonotopy in human auditory cortex examined with functional magnetic resonance imaging. *Hum. Brain Mapp.* 5:18-25.
- Whalen, D. H., 1991. Subcategorical phonetic mismatches and lexical access. *Percept. Psychophys.* 50:351-360.
- Wise, R., F. Chollet, U. Hadar, K. Friston, E. Hoffner, and R. Frackowiak, 1991. Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain* 114:1803-1817.\*
- Wong, D., R. T. Miyamoto, D. B. Pisoni, M. Sehgal, and G. D. Hutchins, 1999. PET imaging of cochlear-implant and normal-hearing subjects listening to speech and nonspeech. *Hear. Res.* 132:34-42.\*
- Wong, D., D. B. Pisoni, J. Learn, J. T. Gandour, R. T. Miyamoto, and G. D. Hutchins, 2002. PET imaging of differential cortical activation by monaural speech and nonspeech stimuli. *Hear. Res.* 166:9-23.\*
- Yvert, B., A. Crouzeix, O. Bertrand, A. Seither-Preisler, and C. Pantev, 2001. Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cereb. Cortex* 11:411-423.
- Zwitserslood, P., 1989. The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32:25-64.
- Zwitserslood, P., and H. Schriefers, 1995. Effects of sensory information and processing time in spoken-word recognition. *Lang. Cogn. Process.* 10:121-136.

Figure 1 (p. 33). Meta-analysis results for passive listening experiments with silent control conditions. Colors indicate the areas of reliable overlap of activation foci for the different stimuli. Results are projected onto sagittal slices of the temporal lobe at Talairach x-coordinates -51 (upper panel) and +51 (lower panel). To facilitate orientation, the approximate contours of the 50-75 % probability volumes of the primary auditory cortices (Penhune et al., 1996) at  $x = -51$  and  $x = +51$  are indicated by a black line. (SY = sylvian fissure, STS = superior temporal sulcus, ITS = inferior temporal sulcus)

Figure 2 (p. 34). Synopsis of hemodynamic activation foci reported for passive listening experiments with silent control conditions. Numbers refer to the experiments listed in Table 1. To facilitate orientation, the approximate contours of the 50-75 % probability volumes of the primary auditory cortices (Penhune et al., 1996) at  $x = -51$  and  $x = +51$  are indicated by a black line. (SY = sylvian fissure, STS = superior temporal sulcus, ITS = inferior temporal sulcus)

Figure 3 (p. 35). Synopsis of hemodynamic activation foci reported for passive listening experiments with heterogeneous auditory control conditions. Numbers refer to the experiments listed in Table 2. To facilitate orientation, the approximate contours of the 50-75 % probability volumes of the primary auditory cortices (Penhune et al., 1996) at  $x = -51$  and  $x = +51$  are indicated by a black line. (SY = sylvian fissure, STS = superior temporal sulcus, ITS = inferior temporal sulcus)

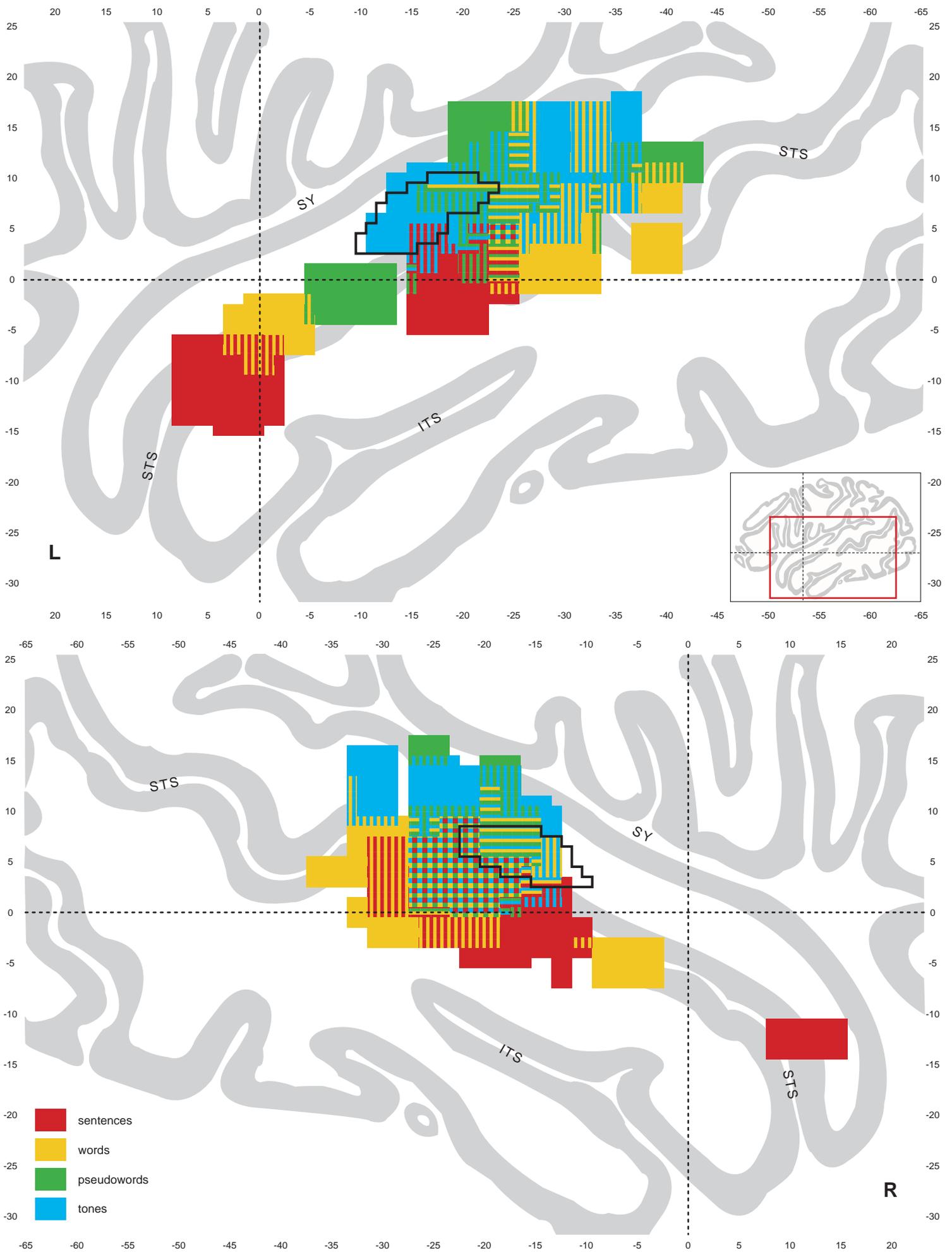


Figure 1

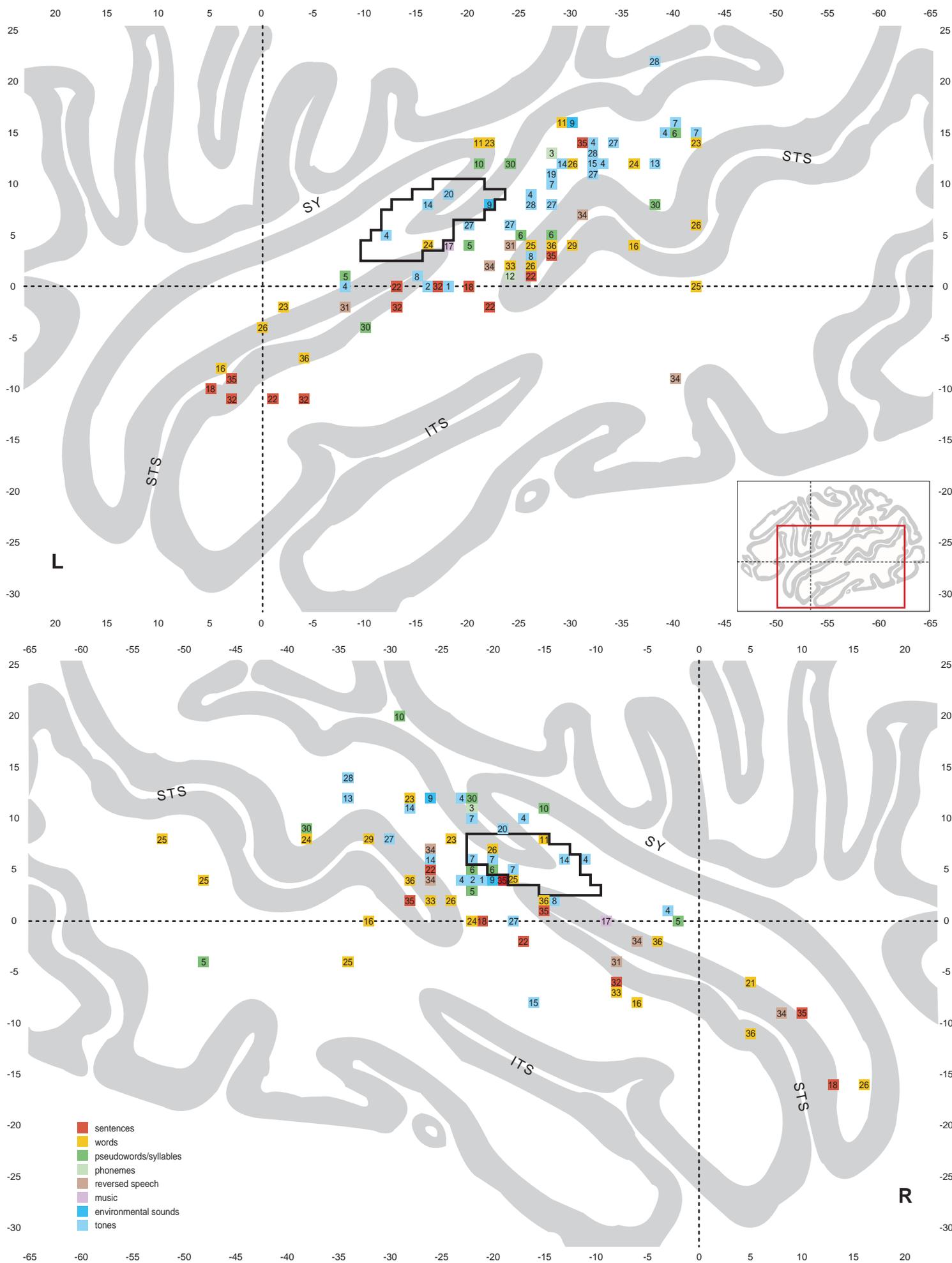


Figure 2

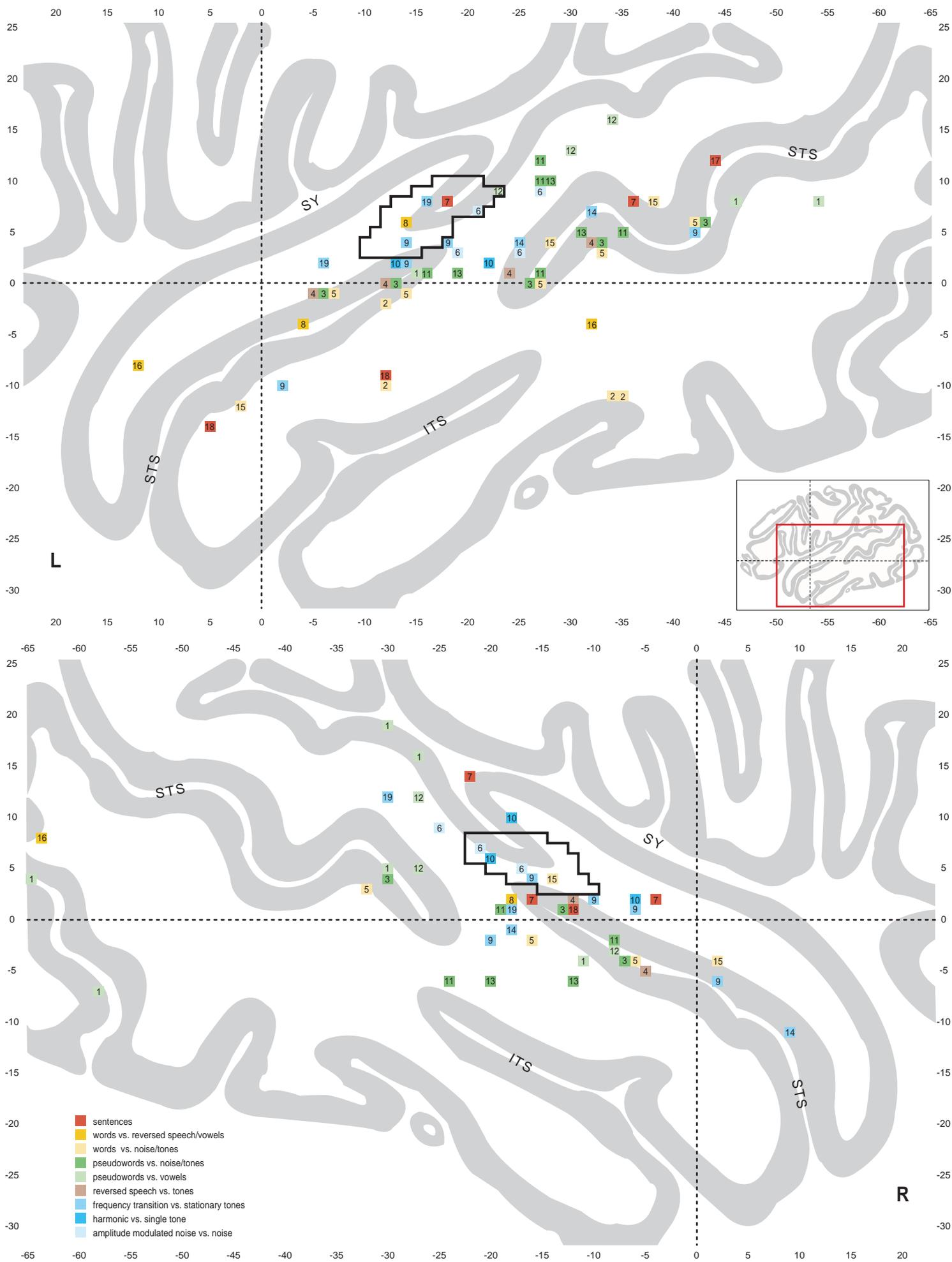


Figure 3