Processing Lexical Semantic and Syntactic Information in First and Second Language: fMRI Evidence From German and Russian

Shirley-Ann Rüschemeyer,¹ Christian J. Fiebach,^{1,2} Vera Kempe,³
and Angela D. Friederici¹

¹Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

²Department of Psychology and Hellen Wills Institute of Neuroscience, University of California, Berkeley, California, USA

³Department of Psychology, University of Stirling, Stirling, Scotland, United Kingdom

Abstract: We introduce two experiments that explored syntactic and semantic processing of spoken sentences by native and non-native speakers. In the first experiment, the neural substrates corresponding to detection of syntactic and semantic violations were determined in native speakers of two typologically different languages using functional magnetic resonance imaging (fMRI). The results show that the underlying neural response of participants to stimuli across different native languages is quite similar. In the second experiment, we investigated how non-native speakers of a language process the same stimuli presented in the first experiment. First, the results show a more similar pattern of increased activation between native and non-native speakers in response to semantic violations than to syntactic violations. Second, the non-native speakers were observed to employ specific portions of the frontotemporal language network differently from those employed by native speakers. These regions included the inferior frontal gyrus (IFG), superior temporal gyrus (STG), and subcortical structures of the basal ganglia. *Hum Brain Mapp* 25:266–286, 2005. © 2005 Wiley-Liss, Inc.

Key words: lexical semantic information; syntactic information; fMRI; inferior frontal gyrus; superior temporal gyrus; basal ganglia

INTRODUCTION

Spoken language comprehension is dependent on the successful breakdown, analysis, and (re)integration of information by the listener. In all natural human languages, this information is encoded on a number of linguistic levels, e.g., phonology, prosody, semantics, and syntax, all of which are processed on the temporal order of milliseconds and are in

some manner represented within a frontotemporal network in the human brain. We compare the neural networks underlying semantic and syntax processing in native speakers of two different languages, namely Russian and German. We then compare brain activation brought on by German as a native language with activation brought on by German as a foreign language (learned by Russian natives). Differences in brain activation between the two groups of native speakers should only be minor under the assumption that a universal language network, which is not language specific, underlies the human capacity to process language in general. In contrast, substantial differences are expected for the comparison of changes in the hemodynamic response elicited by native versus non-native speakers of German, each presented with German sentences. Such differences have been captured previously in electrophysiologic studies [Hahne, 2001].

*Correspondence to: Shirley-Ann Rüschemeyer, Max-Planck-Institute for Human Cognitive and Brain Sciences, Postfach 500 355, D-04303 Leipzig, Germany. E-mail: ruesch@cbs.mpg.de

Received for publication 19 March 2004; Accepted 15 October 2004

DOI: 10.1002/hbm.20098

Published online in Wiley InterScience (www.interscience.wiley.com).

For native speakers, distinct event-related potential (ERPs) have been shown to correlate with different aspects of sentence comprehension: Phonological categorization and phoneme recognition have been postulated in different early ERP components around 150–200 ms after presentation of a word [Connolly and Phillips, 1994; Näätänen et al., 1997], and semantic processing has been related to a negative wave component peaking approximately 400 ms after word presentation (N400) [Kutas and Van Petten, 1994]. Syntactic processing has been postulated to be reflected in a biphasic ERP pattern comprising an early, automatic word category decision approximately 150 ms after word onset (ELAN) [Friederici et al., 1996; Friederici, 2002] and a second, later component peaking around 600 ms after word onset, thought to reflect processes of final syntactic integration (P600) [Friederici, 2002; Kaan et al., 2000; Osterhout et al., 1994].

Importantly, these language-related ERP components are not specific to any one language; rather, they have been observed using stimuli from many different languages, including English [Kutas and Van Petten, 1994], German [Hahne and Friederici, 1999], Dutch [Hagoort and Brown, 2000], Japanese [Nakagome et al., 2001], Hebrew [Deutsch and Bentin, 2001], and Italian [Angrilli et al., 2002]. Clearly, specific components are elicited as a result of specific modification of language stimuli (i.e., syntactic manipulation or semantic manipulation) and therefore not all components are found in each study. All components, however, have been observed in studies using language stimuli in a variety of different languages, and thus these ERP signatures are not in any way tied to a specific language.

Neuroimaging methods have been used increasingly in recent years to investigate language processing; however, a fully satisfactory neuroanatomic model incorporating the many facets of language processing has not yet emerged. Although many studies show overlapping sites of activation correlating with specific aspects of language processing [see Abutalebi et al., 2001, for a review], there are also discrepancies. Differences may be attributed to the use of different sentence types, presentation modalities, and tasks; therefore, only a preliminary functional description of different brain regions is currently possible. We outline briefly some of the more consistent brain activations reported for various specific aspects of language processing.

The importance of perisylvian cortex in language processing has been clear for several decades based on lesion studies and electrocortical stimulation. In vivo data obtained from healthy individuals with newer neuroimaging techniques such as functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) have helped to identify correlations between increased activation in specific cortical areas and the processing of specific linguistic functions [Price, 2000]. On the single-word processing level, the left angular gyrus and left temporal lobe, particularly the middle and inferior temporal gyri, have been shown to be involved in long-term storage of semantic information [Price, 2000]. Left inferior frontal cortex, in par-

ticular Brodmann's areas (BA) 45 and 47, have been postulated to support the retrieval (not storage) of semantic information and the processing of semantic relationships between words [Bookheimer, 2002]. At the sentence level, processing of syntactic structure has been shown to activate anterior regions within the superior temporal gyrus (STG) [Friederici et al., 2003; Humphries et al., 2001; Meyer et al., 2000]. Additionally, portions of inferior frontal cortex (BA44) are activated increasingly as an effect of syntactic complexity [Caplan et al., 1998, 1999; Just et al., 1996] and working memory demands within sentences (longer filler-gap dependencies) [Cooke et al., 2001; Fiebach et al., 2001, 2004].

Several studies have directly compared syntactic with semantic processing in auditory sentence comprehension [Dapretto and Bookheimer, 1999; Friederici et al., 2000, 2003; Kuperberg et al., 2000; Meyer et al., 2000; Ni et al., 2000]. A general trend in these studies points to a greater involvement of temporal cortex in sentential semantic processing versus syntactic processing (although portions of STG have been implicated in syntax processing as outlined above) [Friederici et al., 2003; Kuperberg et al., 2000; Ni et al., 2000] and a distribution of functional specialty tied to semantic and syntax processing within frontal cortex [Dapretto and Bookheimer, 1999; Friederici et al., 2000, 2003; Meyer et al., 2000; Ni et al., 2000]. Specifically, within frontal cortex it has been suggested that anterior portions of inferior frontal gyrus (IFG; BA45/47) are engaged in semantic processing, whereas posterior portions (BA44/45) are responsible specifically for the processing of syntax in sentences [Bookheimer, 2002; Dapretto and Bookheimer, 1999]. This has led to the proposal that two separate temporofrontal networks in the left hemisphere support semantic and syntactic processes respectively [Friederici, 2002; Friederici and Kotz, 2003].

Again, these neuroanatomic substrates are not assumed to be unique to any one spoken language, although the processing of specific surface features of different languages (i.e., orthography) has been argued to have an effect on processing strategies [Paulesu et al., 2000; Tan et al., 2003]. Activations in the aforementioned classical language areas as a function of higher-level linguistic processing (i.e., processing of semantic vs. syntactic features) have been observed not only for English and German, but also for Italian [Moro et al., 2001] and Japanese [Suzuki and Sakai, 2003].

The present study set out to investigate the assumption that the neural basis of semantic and syntactic process is universal by directly comparing the processing of syntactic and semantic violations in two languages with quite different underlying syntactic structures (Experiment 1). To this end, German sentences were presented to native German participants, and Russian sentence stimuli were presented to a group of native Russian participants. In each case, the stimulus materials used elicited reliable effects at the electrophysiological level [Hahne, 2001]. Using fMRI, we recorded changes in the hemodynamic response of participants to examine the brain regions that are involved in sentence

processing in two different native languages. Given the similar ERP patterns seen in native speakers of different languages in response to language stimuli in their native language, we expected to see similar brain regions activated in the processing of different languages by native speakers.

The second question addressed in this study (Experiment 2) pertained to differences in the processing of a native (L1) versus a foreign (L2) language. To this end, we compared the data obtained for native German speakers in Experiment 1 with results of non-native speakers listening to the same sentences. Again, electrophysiologic results have shown reliably that differences do exist, at least temporally, between the processing of L1 and L2 [Hahne, 2001]. Early anterior negativities in ERP responses to word category and morpho-syntactic violations are usually absent in non-native speakers, and later integrative components related to semantic and syntactic structure violations typically show a reduction in amplitude as well as a shift in latency [for review, see Mueller, in press].

Previous neuroimaging studies have provided heterogeneous results regarding the processing of a second language. Some studies argue for similar cortical networks supporting L1 and L2 processing [Chee et al., 1999a,b, 2000; Kim et al., 1997; Luke et al., 2002; Nakada et al., 2001; Perani et al., 1998; Tan et al., 2003] whereas others argue for differential processing networks [Dehaene et al., 1997; Perani et al., 1996, 2003; Wartenburger et al., 2003; Yetkin et al., 1996]. In those studies arguing for common cortical networks, it has been shown that different portions of the same network may be employed differentially for native and non-native speakers; however, the localization of processing areas remains common to both [Chee et al., 1999b; Kim et al., 1997]. Differences in materials, methods, and modalities certainly play a role in these discrepancies. For example, those studies investigating L2 word generation in homogenous groups of bilinguals tend to show common neural responses for L1 and L2, pointing to a shared mental lexicon containing conceptual information used by both language systems [Chee et al., 1999b; Yetkin et al., 1996]. A recent study by Wartenburger et al. [2003], however, demonstrates that the cerebral organization underlying semantic processing systems is influenced heavily by the bilingual's proficiency level, thus pointing to a different organization in less-proficient bilinguals. Because so many variables exist in relation to the participants studied and the linguistic materials, a clear line simply does not exist in the literature. Many studies have shown, however, that age of L2 acquisition, proficiency level, and exposure to a second language all influence language-processing strategies [Kim et al., 1997; Perani et al., 1998, 2003; Wartenburger et al., 2003; Yetkin et al., 1996]. Furthermore, it seems generally true that less fluency in a language is characterized by more variability in cortical areas supporting the processing of that language, a fact that has been shown to confound the results of studies looking at group averages of second-language users [Dehaene et al., 1997; Yetkin et al., 1996].

We wished to examine whether the neural substrates supporting L1 and L2 processing are the same or different if the syntactic structure of the presented linguistic stimuli is present in each group's native language. Based on the results of previous studies, we attempted to keep our participant group as homogenous as possible, selecting highly proficient non-native speakers of German to take part in experimental testing. We also explored whether or not differential neural correlates for syntactic versus semantic processing systems could be targeted in non-native speakers, and to what extent such specific processing systems might influence language processing as a whole.

SUBJECTS AND METHODS

Experiment 1: Syntactic and Semantic Processes in Native Speakers

This experiment presents data from native German speakers processing German sentences and native Russian speakers processing Russian sentences.

Participants

After giving informed consent, 18 native speakers of German (8 men; mean age, 25 years; age range, 23–30 years) and 7 native speakers of Russian (3 men; mean age, 30.5 years; age range, 23–32 years) participated in the study. The results of the original group of German participants are reported elsewhere [Friederici et al., 2003]. To make the groups of native speakers more comparable in terms of size, we randomly selected 7 German native speakers from the original 18 for further analysis. The Russian native speakers investigated in this study were second-language learners of German, and had been living in Germany for an average of 7 years. No participant had any history of neurologic or psychiatric disorders. All participants had normal or corrected to normal vision, and were right-handed (laterality quotients of 90–100 according to the Edinburgh handedness scale) [Oldfield, 1971].

Materials

German. The experimental material consisted of short sentences containing transitive verbs in the imperfect passive form. Participial forms of 96 different transitive verbs, all of which started with the regular German participial morpheme *ge*, were used to create the experimental sentences. For each participle, three different critical sentences and one filler sentence were constructed (see Table I).

All sentences began with a noun phrase made up of a definite article (*der*, *die*, or *das*) and an uninflected singular noun. This noun phrase was followed by the imperfect form of the passive auxiliary verb *werden*. At this point, sentences in all conditions were constructed identically. In the correct sentences, the participial form of a transitive verb directly followed the auxiliary verb, thus creating a short, acceptable sentence in the imperfect passive tense. In the syntactically incorrect sentences, the auxiliary verb was followed directly

TABLE I. Sentence materials

Sentence	German	Russian
COR	Das Brot wurde <u>gegessen</u> .	Ja dumaju, chto produkty <u>prinesut</u> .
	The bread was eaten.	<i>I think that the food is brought.</i>
SYN	Das Eis wurde im <u>gegessen</u> .	Ja dumaju, chto ovoschi dlya <u>prinesut</u> .
	<i>The ice-cream was in-the eaten.</i>	<i>I think that the vegetables for-the are brought.</i>
SEM	Der Vulkan wurde <u>gegessen</u> .	Ja dumaju, chto grom <u>prinesut</u> .
	<i>The volcano was eaten.</i>	<i>I think that the thun der is brought.</i>

Examples of sentence materials (COR, correct sentences; SYN, syntactically violated sentences; SEM, semantically violated sentences) presented in German (Experiment 1 and 2) and Russian (Experiment 1), plus English translation equivalents. The critical word in each sentence is underlined. English translations maintain their original word order.

by an inflected preposition, which suggests the initiation of a second noun phrase. The presence of a second noun phrase at this position is entirely acceptable in German, thus the preposition alone poses no problem. This preposition, however, must be followed by the remaining missing elements of the noun phrase, most critically by a noun. Precisely this was violated in the syntactically incorrect sentences: Immediately after the preposition, the sentence final verb participle was presented instead of the necessary noun. This yielded a clear phrase structure violation. The inflected forms of seven prepositions (*in*, *zu*, *unter*, *vor*, *am*, *bei*, and *für*) were used to construct the syntactically incorrect sentences. Semantically incongruous sentences had the same grammatical form as correct sentences did (noun phrase followed by verb phrase consisting of an auxiliary and the participle form of a transitive verb); however, the lexical-conceptual meaning of the participle could not be integrated satisfactorily with the preceding sentence context. One final condition, which was not included in the final fMRI analysis, was also presented. This condition constituted correct sentences with the form noun phrase, followed by the auxiliary, followed by a completed prepositional phrase (preposition and necessary noun), followed finally by the participle form of the verb. This filler condition served two purposes: it allowed the number of correct and incorrect sentences to be balanced and it prevented participants from being able to determine the grammaticality of sentences based solely on the presence of a preposition. In other words, the mere presence of a preposition was not sufficient for predicting sentence acceptability.

The sentences were spoken by a trained female native speaker, recorded and digitized, and presented acoustically to the participants. The complete set of materials is available from the authors.

Russian. We attempted to create Russian sentences that were as similar as possible in terms of their syntactic structure to the German sentences described above. This required us to take into consideration a number of features that are specific to the Russian language and could potentially increase variability in the experimental sentences. These constraints are discussed below.

Many Russian prepositions are homonymous to verb prefixes. This may render the targeted syntactic violation ambiguous because in acoustically presented stimuli, the prepositions may be interpreted as verb prefixes and therefore may be taken as legal continuations of the sentence. Only later, when a mismatch between the prefix and the verb occurs, can a violation be detected. Such a violation, however, would be morphologic in nature and thus completely different from the intended syntactic violation. To avoid this ambiguity, we attempted to use prepositions that never occur as verb prefixes. However, pilot studies revealed that some of these prepositions, namely *cherez* (spatially: through, over; temporally: in), *okolo* (by, next to), *pered* (spatially: in front of; temporally: before), and *vozle* (by, next to) failed to elicit the ERP patterns typical for syntactic violations. Only the preposition *dlya* (for) was suitable to create syntactic violations, so this preposition was used for the entire set of stimulus materials.

For the semantic and syntactic violation to occupy identical positions, the main verb had to be sentence final. This allowed us to keep word order as similar as possible across all sentence types. Moreover, to minimize differences in the prosodic contour, the sentence-final verb received stress in all sentences. Russian permits considerable freedom in word order, and stress is used to indicate the focus of the sentence. Sentence-final stress indicates either contrastive or non-contrastive focus on the verb, whereas stress elsewhere always indicates contrastive focus. Sentence-final stress was thus the only possibility to keep the sentences identical with respect to prosodic contour and informational structure. It also eliminated any early prosodic cues that would give away the upcoming sentence structure. Care was taken to ensure that the minimal focus on the verb imposed by sentence-final stress would not conflict with the meaning of the whole sentence.

Russian verbs can be either intransitive (unaccusative or unergative) or passivized transitive. It was necessary to use both verb types, because it proved impossible to find 160 different subjects of intransitive verbs. The factor of verb type, however, was counterbalanced by constructing half of the sentences with intransitive verbs and half with passivized transitive verbs.

These unavoidable differences in verb type resulted in different syntactic functions of the noun phrase (NP). In sentences with intransitive verbs, the NP was the subject of the sentence, whereas in sentences with passivized transitive verbs, the NP was the object of the sentence. (Note that Russian permits null-subject sentences.) In Russian, the syntactic function of an NP can also be marked by its morphologic case. Nominative case marks the subject and accusative

case marks the object of a sentence; however, nominative-accusative marking is often neutralized in Russian nouns, so that the morphologic structure of the noun sometimes does and sometimes doesn't provide unambiguous information about the function of the NP. Specifically, regular feminine singular and animate masculine singular nouns provide unambiguous subject and object marking. All other nouns (inanimate masculine, irregular feminine, neuter, and plural) do not. To control for case-marking ambiguity, feminine and masculine animate nouns were avoided. The NP was thus always morphologically ambiguous with respect to case marking, and only the verb indicated whether the NP was subject or object of the sentence.

To avoid confounds associated with coarticulation and prosodic cues on the preposition *dlya*, all sentences in the syntactic violation condition were recorded inserting bisyllabic nonwords after the preposition. The nonwords were composed of the first syllable of the upcoming verb and the preposition *dlya*. For example, the sentence *Ja nadjus', chto polotence dlya vysochnet* (I hope that the towel for will dry), which contains a syntactic violation, was recorded as *Ja nadjus', chto polotence dlya vydlja vysochnet*. During recording, the speaker attempted to produce a prosodic pattern of the sentence such as if the nonword *vydlja* was a noun. All sentences were subsequently digitized, and the nonword was deleted out of sentences containing a syntactic violation. The spliced sentences do not sound acoustically unusual to native speakers. Using the above-mentioned constraints, we constructed 160 sentences in total with 40 sentences in each condition.

Experimental procedure

German. Two differently randomly ordered stimulus sequences were designed for the experiment. The 96 sentences from each of the four conditions were distributed systematically between two lists, so that each verb occurred in only two of four conditions in the same list. Forty-eight null events, in which no stimulus was presented, were also added to each list. The lists were then pseudorandomly sorted with the constraints that: (1) repetitions of the same participle were separated by at least 20 intervening trials; (2) no more than three consecutive sentences belonged to the same condition; and (3) no more than four consecutive trials contained either correct or incorrect sentences. Furthermore, the regularity with which two conditions followed one another was matched for all combinations. The order of stimuli in each of the two randomly sorted stimulus sequences was then reversed, yielding four different lists. These were distributed randomly across participants.

An experimental session consisted of three 11-min blocks. Blocks consisted of an equal number of trials and a matched number of items from each condition. Each session contained 240 critical trials, made up of 48 items from each of the four experimental conditions plus an equal number of null trials, in which no stimulus was presented and blood oxygenation-dependent (BOLD) response was allowed to return to a baseline state.

The 240 presented trials lasted 8 s each (i.e., four scans of repetition time [TR] = 2 s). The onset of each stimulus presentation relative to the beginning of the first of the four scans was varied randomly in four time steps (0, 400, 800, and 1,200 ms). The purpose of this jitter was to allow for measurements to be taken at numerous time points along the BOLD signal curve, thus providing a higher resolution of the BOLD response [Miezin et al., 2000]. After the initial jittering time, a fixation cue consisting of an asterisk in the center of the screen was presented for 400 ms before presentation of the sentence began. Immediately after hearing the sentence, the asterisk was replaced by three question marks, which cued participants to make a judgment on the correctness of the sentence. Maximal response time allowed was 2,000 ms. Identifying the type of error was irrelevant. Participants indicated their response by pressing buttons on a response box, and after the response the screen was cleared. Incorrect responses and unanswered trials elicited a visual feedback. These trials, as well as two dummy trials at the beginning of each block, were not included in the data analysis.

Russian. Two different randomly sorted stimulus sequences were designed for Russian sentences as well. The 40 sentences from each condition were ordered pseudorandomly with the constraints that: (1) repetitions of a participle and null-events never occurred; (2) no more than three consecutive sentences belonged to the same condition; and (3) no more than four consecutive trials contained either correct or incorrect sentences. The regularity with which two conditions followed one another was matched for all combinations. The order of stimulus in each of the two randomly sorted stimulus sequences was reversed, yielding four different lists. These were distributed randomly across participants.

An experimental session consisted of three 11-min blocks. Blocks consisted of an equal number of trials and a matched number of items from each condition. Each session contained 200 critical trials, made up of 40 items from each of the four experimental conditions plus an equal number of null trials in which no stimulus was presented (see above).

The 200 presented trials lasted 10 s each (i.e., five scans of TR = 2 s). Trials were made 2 s longer in comparison to the German sentences to allow better the BOLD response to return to baseline. The onset of each stimulus presentation relative to the beginning of the first of the five scans was varied randomly between 0, 500, 1,000 or 1,500 ms. Again, this parameter differs from that used for the presentation of German sentences, where a jitter of 0, 400, 800, or 1,200 ms was used. We determined 500 ms to be a more intuitively logical jittering step, but did not anticipate that this would cause any great difference in the data obtained for Germans and Russians. The presentation procedure was in all other respects identical to that for German sentences (see above).

Functional MRI data acquisition

In the first group of German native participants, eight axial slices (5 mm thickness, 2 mm interslice distance, field of

view [FOV] 19.2 cm, data matrix of 64×64 voxels, and in-plane resolution of $3 \text{ mm} \times 3 \text{ mm}$) were acquired every 2 s during functional measurements (BOLD-sensitive gradient echo-planar imaging [EPI] sequence, TR = 2 s, echo time [TE] = 30 ms, flip angle 90 degrees, and acquisition bandwidth 100 kHz) with a 3-Tesla Bruker Medspec 30/100 system. In the group of Russian native speakers, 10 axial slices of the same dimensions were obtained. Before functional imaging, T1-weighted modified driven equilibrium Fourier transform (MDEFT) images (data matrix of 256×256 , TR = 1.3 s, and TE = 10 ms) were obtained with a nonslice-selective inversion pulse followed by a single excitation of each slice [Norris, 2000]. These were used to coregister functional scans with previously obtained high-resolution whole-head 3-D brain scans (128 sagittal slices, 1.5 mm thickness, FOV $25.0 \times 25.0 \times 19.2 \text{ cm}$, and data matrix of 256×256 voxels) [Lee et al., 1995].

Data analysis

The functional imaging data processing was carried out using the software package LIPSIA [Lohmann et al., 2001]. Functional data were corrected first for motion artifacts and then for slicetime acquisition differences using sinc-interpolation. Low-frequency signal changes and baseline-drifts were removed by applying a temporal high-pass filter to remove frequencies lower than 1/60 Hz. A spatial filter of 5.65 mm full-width half-maximum (FWHM) was applied. The anatomic images acquired during the functional session were coregistered with the high-resolution full-brain scan and then transformed by linear scaling to a standard size [Talairach and Tournoux, 1988]. This linear normalization process was improved by a subsequent processing step that carried out an additional nonlinear normalization [Thirion, 1998]. The transformation parameters obtained from both normalization steps were subsequently applied to the preprocessed functional images. Voxel size was interpolated during coregistration from $3 \text{ mm} \times 3 \text{ mm} \times 5 \text{ mm}$ to $3 \text{ mm} \times 3 \text{ mm} \times 3 \text{ mm}$. Statistical evaluation was based on a least-squares estimation using the general linear model for serially autocorrelated observations [Worsley and Friston, 1995]. The design matrix was generated with a synthetic hemodynamic response function (HRF) [Friston et al., 1998; Josephs et al., 1997]. The model equation, made up of the observed data, the design matrix, and the error term, was convolved with a Gaussian kernel of dispersion of 4 s FWHM. For each participant, two contrast images were generated, which represented the main effects of: (1) syntactically violated sentences versus correct sentences; and (2) semantically violated sentences versus correct sentences. The group analysis consisted of a one-sample *t*-test across the contrast images of all subjects that indicated whether observed differences between conditions were significantly distinct from zero, and *t*-values were transformed subsequently into Z-scores. The resulting *t*-statistics were transformed to standard normalized distribution. Group statistical parametric maps (SPM[Z]) were thresholded at $Z > 2.57$

($P < 0.005$, uncorrected). Only clusters of at least 14 connected voxels (i.e., 400 mm^3) were reported.

Penetrance maps evaluating the consistency of group results across participants were calculated as outlined by Fox et al. [1996]. Z-images from each participant characterizing differences in activation between syntactic errors and correct sentences between semantic errors and correct sentences were converted to binary images (each voxel valued at either 0 or 1), based on a Z-threshold of $P < 0.1$. The binary images were then summed across the 7 participants. Resulting maps are color-coded representations of the number of participants showing significant differences in activation in each voxel.

To confirm the validity of the statistical differences observed in direct contrasts, those areas showing an increase in mean signal change were subjected to a subsequent region-of-interest (ROI) analysis. Mean activation from the peak voxel determined in the direct contrasts was calculated for each participant over a 5-s timeframe (3–8 s after presentation of the critical word). These values were then used in a repeated-measures analysis of variance (ANOVA) of mean signal change.

Experiment 2: Processing Native Versus Foreign Language

Experiment 2 was conducted with the goal of comparing brain activation patterns for natives and non-native speakers in a sentence comprehension task.

Participants

After giving informed consent, 18 native speakers of German (the same participants from Experiment 1) and 14 non-native speakers of German (3 men; mean age, 25.6 years; age range, 22–30 years) participated in the study. Non-native German speakers were native speakers of Russian, and had been living in Germany in for an average of 5 years. Of the 14 non-native participants, 6 were also participants in Experiment 1. No participant had any history of neurologic or psychiatric disorders. All participants had normal or corrected to normal vision, and were right-handed (laterality quotients of 90–100 according to the Edinburgh handedness scale) [Oldfield, 1971].

Material and experimental procedure

The same German materials and experimental procedure were used as for the German native speakers in Experiment 1.

Functional MRI data acquisition and analysis

The data was obtained in an identical manner (i.e., same scanner and identical parameters) to that used for German natives from Experiment 1.

A within-group analysis of each participant group was made in an identical manner to that described in Experiment

TABLE II. Reaction times and accuracy for native speakers

Language	Reaction time (ms)			Accuracy (%)		
	COR	SEM	SYN	COR	SEM	SYN
German	417 ± 34	449 ± 36	418 ± 33	97 ± 1.0	92 ± 3.7	97 ± 1.3
Russian	417 ± 43	406 ± 53	394 ± 43	92 ± 1.6	93 ± 1.7	94 ± 2.3

Reaction times (ms) and accuracy (all values given as mean ± standard error) for participants listening to correct sentences (COR), semantically anomalous sentences (SEM) and syntactically anomalous sentences (SYN) in their respective native languages.

1, with the following exceptions. SPMs were thresholded at $Z > 3.09$ ($P < 0.001$, uncorrected). Only clusters of at least 14 connected voxels (400 mm^3) were reported. We were able to use a more stringent threshold in Experiment 2 than in Experiment 1, as we collected data from twice as many participants. In Experiment 1, we attempted to compensate for the low number of participants by showing the results of penetrance maps. In Experiment 2, we obtained data from a large group of participants in each group, and were therefore able to show results at a higher statistical threshold.

For between-group comparisons, two-sample t -tests were conducted comparing contrast images from individuals in each group (group of native speakers vs. group of non-native speakers) while listening to each experimental sentence type against a resting baseline. High levels of variance, in particular within the group of non-native speakers, made the detection of stable effects in the third-level analysis between the two groups of participants more difficult than were direct contrasts observed in the second-level analysis within the groups. We therefore followed other researchers in lowering threshold levels to $Z > 2.32$ ($P < 0.01$) in determining the significance of between-group differences [Pallier et al., 2003; Perani et al., 1998]. Furthermore, to determine that the size of observed activations was reliably different between groups, we conducted a second third-level analysis based on Bayesian statistics, which provided a probability estimate for the reliability of difference in activation size (expressed as a percentage value between 0 and 100) and was not susceptible to problems of multiple comparisons [Neumann and Lohmann, 2003]. To do this, the peak coordinate obtained from the direct contrasts was tested in each group of participants. We reported Bayesian statistics for the between-group comparisons only, as we wished to solidify the statistical significance of our results in these contrasts. In within-group contrasts presented in both Experiments 1 and 2, this additional statistical exploration was not necessary.

To confirm further the validity of the statistical differences observed in direct contrasts, those areas showing an increase in mean signal change were subjected to a subsequent ROI analysis. Mean activation from the peak voxel determined in the direct contrasts was calculated for each participant over a 6-s timeframe. Mean signal change over this timeframe was then used in for repeated-measures ANOVA.

RESULTS

Experiment I

Behavioral results

Reaction times. Repeated-measures ANOVAs were calculated for the two groups (German natives and Russian natives) and the three experimental conditions (correct sentences [COR], syntactically anomalous sentences [SYN], and semantically anomalous sentences [SEM]). Only trials that were answered correctly were included in the analysis. Furthermore, trials that had a reaction time deviating from the group average by 2.5 standard deviations (SD) or more were excluded. No significant main effect of group or condition could be observed. Reaction times in this study do not reflect online sentence processing because participants were asked to wait for a cue before answering (Table II).

Error rates Repeated-measures ANOVA for the groups described above were calculated. The results yielded no significant main effects and no interaction between group and condition (Table II).

Imaging results

Talairach coordinates for the activations discussed here can be found in Table III. Images of selected activations, penetrance maps depicting stability of activations, and time-courses showing the percent signal change for each condition over the course of a trial are shown in Figure 1. Mean ROI values are plotted in Figure 2.

Direct contrasts

Syntactic processes were investigated in a direct comparison of syntactically violated sentences versus correct sentences. This comparison showed more activation for syntactic anomalies than for correct sentences within the mid-STG in each group of participants listening to their respective native language. In both groups, this activation was lateral to Heschl's gyrus and extended into cortex slightly anterior to the primary auditory cortex. Greater activation levels were observed for correct sentences in comparison to that for syntactically anomalous sentences in the posterior cingulate and inferior precuneus region. Posterior cingulate activation was observed for correct sentences in both groups

TABLE III. Talairach coordinates, Z-values, and volume of the activated regions for the different contrasts in native speakers

Contrast	<i>x</i>	<i>y</i>	<i>z</i>	Z-max	Volume	Region
German						
SYN-COR	-56	-19	12	3.90	752	L STG, maximum and posterior peak
	-55	-5	11	3.55		L STG, lateral anterior peak
COR-SYN	-43	-58	35	3.11	538	L Posterior STS, ascending branch
	4	-49	32	2.96		R Precuneus
SEM-COR	-43	20	6	3.93	1,539	L IFG
COR-SEM	-13	-55	35	3.85	1,995	L Precuneus
	7	-43	21	3.33		R Posterior cingulate gyrus
Russian						
SYN-COR	-47	-28	9	3.81	2,091	L STG, maximum
	-59	-21	12	3.29		L STG, lateral posterior peak
	-58	-5	8	3.18		L STG, lateral anterior peak
COR-SYN	-10	-52	30	3.88	6,440	L Precuneus
SEM-COR	-49	26	-6	4.02	822	L IFG, pars orbitalis (BA47)
COR-SEM	-4	-43	44	3.40	534	L Posterior cingulate

Talairach coordinates (*x*, *y*, *z*), Z-values and volume (mm³) of the activated regions for the different contrasts: syntactically anomalous sentences vs. correct sentences (SYN-COR), semantically anomalous sentences vs. correct sentences (SEM-COR), and correct sentences vs. each anomalous condition (COR-SYN, COR-SEM). Z-values were thresholded at $Z > 2.57$ ($P < 0.005$, uncorrected) and clusters had a minimum size of 14 voxels (400 mm³). L, left; R, right; STG, superior temporal gyrus; IFG, inferior frontal gyrus.

of participants. Analysis of the time-courses obtained from this region, however, showed that activation differences were not a reflection of an increase in signal change in response to correct sentences, but rather a decrease in signal change in response to the anomalous condition.

Semantic processes were focused on in a direct comparison of semantically violated sentences versus correct sentences. This comparison revealed increased levels of activation in IFG in both groups of participants irrespective of native language in response to semantic anomalies. The peak of this activation lay within the pars orbitalis of the IFG (BA45/47) in both groups. Differential activation was again observed in the left posterior cingulate gyrus and precuneus region, with a greater activation level in response to correct sentences as compared to that for semantically anomalous sentences. As in the comparison with syntactically incorrect sentences, this pattern did not reflect an increase in signal change for the correct condition and is thus not discussed further.

Penetrance maps

Results from the penetrance maps indicating the number of participants showing significant difference in activation between conditions show that reported group differences were relatively stable across participants. Maps of all contrasts except one show good consistency between group averages and individual activation patterns.

ROI analysis

Two critical ROIs were defined for this experiment: the left STG centered around the more anterior peak activation observed in both groups ($-58, -5, 8$) and the left IFG ($-43,$

$20, 6$). After obtaining the mean signal change from each individual participant in each ROI, ANOVAs were calculated with the independent variables ROI (STG and IFG), group (German and Russian), and the repeated factor condition (COR, SEM, and SYN). The two groups of participants showed comparable activation patterns across ROIs and conditions. A three-way interaction between ROI $\times$ group $\times$ condition was not significant; there was likewise no main effect of group. Signal change in the STG was more pronounced than that observed in the IFG, as exemplified by a significant main effect of ROI ($F[1,12] = 23.37$; $P < 0.0005$). Furthermore, the lesser degree of signal change brought on by the condition COR in response to either SEM or SYN in both ROIs led to a main effect of condition ($F[2,24] = 14.96$; $P < 0.0001$).

Across groups, a unique activation pattern for the different conditions within each ROI was observed, as exemplified by the interaction between ROI $\times$ condition ($F[2,24] = 6.81$; $P < 0.005$). Within ROI 1 (left STG), both violation conditions (SEM and SYN) showed more activation than did the correct sentence condition (COR) ($F[2,24] = 11.31$; $P < 0.001$). In ROI 1, the highest level of activation was observed in conjunction with the condition SYN. Within ROI 2 (left IFG), the condition SEM elicited higher levels of activation than did either COR or SYN ($F[2,23] = 11.75$; $P < 0.001$).

Experiment 2

Behavioral results

Reaction times. Repeated-measures ANOVAs were calculated for the two groups (L1 and L2) and the three experimental conditions (correct sentences, syntactically anoma-

Figure 1.

Time-courses (showing percent signal change over time), direct contrast images (showing significance levels over each group of participants), and penetrance maps (showing the number of participants with significant activation increase in each voxel) for native speakers of German and Russian. Values in the direct contrast images indicated by the color bar indicate statistical significance. Values in the penetrance maps refer to numbers of individuals. The upper panel depicts those areas showing increased levels of activation for syntactically incorrect sentences (SYN) in

comparison to correct sentences (COR). Increased activation levels correlating with syntactic violations are seen in left anterior to mid-superior temporal gyrus (STG). The lower panel depicts those areas showing increased levels of activation for semantically incorrect sentences (SEM) in comparison to correct sentences (COR). Increased activation correlating with semantic violations is seen in left inferior frontal gyrus (IFG; BA45/47). Percent signal change in STG is greater than in IFG.

Figure 2.

Experiment 1: Mean percent signal change and standard error for native speakers of German and Russian in each of the ROIs discussed. L STG, left superior temporal gyrus; L IFG, left inferior frontal gyrus.

lous sentences, and semantically anomalous sentences). Only trials that were answered correctly were included in the analysis. Furthermore, trials that deviated from the group average by a factor of 2.5 SD or more were excluded. A main effect of group was observed ($F[1,30] = 9.36$; $P < 0.01$) as was a main effect of condition ($F[2,60] = 13.33$; $P < 0.01$); however, there was no interaction between group and condition. This reflected the fact that the L2 group was slower in responding to sentence stimuli in all conditions, and that both groups were slower in responding to semantically anomalous sentences than to correct sentences (L1, $F[1,17] = 26.75$ and $P < 0.01$; L2, $F[1,13] = 10.27$ and $P < 0.01$) and that the native speakers were slower in responding to semantically anomalous sentences than to syntactically anomalous sentences (L1, $F[1,17] = 20.76$; $P < 0.01$) (Table IV). Only results showing statistical significance after Bonferroni adjustment of the α level were reported. Reaction times as obtained in this experiment were not a reflection of online sentence processing, as participants were told to wait with their judgment selection until prompted.

Error rates. Repeated-measures ANOVAs were calculated with the factors described above. A main effect of group was observed ($F[1,30] = 25.91$; $P < 0.01$) as well as a main effect of condition ($F[2,60] = 7.06$; $P < 0.01$) and a group $\times$ condition interaction ($F[2,60] = 7.92$; $P < 0.01$). Further analysis revealed a significantly greater percentage of errors for L2 speakers than for L1 speakers in all experimental conditions

as well as only a tendency for differences between conditions within the L1 group ($F[2,34] = 2.75$; $P < 0.1$), but a reliable difference between conditions within the L2 group ($F[2,26] = 6.84$; $P < 0.01$). Post-hoc analysis showed that L2 speakers show a tendency to make more errors in the detection of syntactic errors sentences than in judging correct sentences ($F[1,13] = 6.27$; $P < 0.05$) and are significantly better at detecting semantic anomalies than syntactic anomalies ($F[1,13] = 9.39$; $P < 0.01$). The level of significance reported was adjusted according to Bonferroni.

Imaging results

Within-group comparison: non-native speakers. We report direct comparisons between each violation condition and correct sentences for non-native speakers of German. The direct contrasts for native speakers of German are not explicitly elaborated upon here, as a representative subgroup was already discussed in Experiment 1. In addition, the original data is discussed at length elsewhere [Friederici et al., 2003]; however, the coordinates and Z-values of local maxima for the group of native speakers is provided for reference in Table V.

In a direct comparison between syntactically anomalous sentences and correct sentences, non-native speakers showed no areas of differential activation. No areas were seen more activated for syntactically anomalous sentences than for correct sentences; likewise, no areas were seen to be

TABLE IV. Reaction times and accuracy for native and non-native speakers

Language	Reaction time (ms)			Accuracy (%)		
	COR	SEM	SYN	COR	SEM	SYN
L1	372 $\pm$ 18	410 $\pm$ 18	375 $\pm$ 18	97 $\pm$ 0.7	93 $\pm$ 1.7	95 $\pm$ 0.8
L2	489 $\pm$ 35	544 $\pm$ 45	510 $\pm$ 47	85 $\pm$ 2.3	86 $\pm$ 2.2	74 $\pm$ 4.9

Reaction times (ms) and accuracy (all values given as mean $\pm$ standard error) for participants listening to correct sentences (COR), semantically anomalous sentences (SEM) and syntactically anomalous sentences (SYN) in either their native language (L1) or a second language (L2).

TABLE V. Talairach coordinates, Z-values, and volume of the activated regions for different contrasts in native and non-native speakers of German

Contrast	<i>x</i>	<i>y</i>	<i>z</i>	Z-max	Volume	Region
L1						
SYN-COR	-59	-22	12	4.81	2,919	L mid STG
	56	-19	6	4.55	835	R mid STG
COR-SYN	-13	41	15	3.61	1,042	L Superior frontal gyrus
	-5	-43	38	3.71	541	L Posterior cingulate
	-7	-46	24	3.77	561	L Posterior Cingulate
SEM-COR	-40	23	3	5.33	6,082	L IFG (BA45/47)
	41	14	18	4.30	657	R IFG (BA44/6)
	-55	-52	12	4.01	448	L Posterior MTG
COR-SEM	4	-58	47	3.76	878	R Precuneus
	1	-43	30	4.23	4,026	R Posterior cingulate
L2						
SYN-COR	—	—	—	—	—	No significant differences
COR-SYN	—	—	—	—	—	No significant differences
SEM-COR	-53	17	21	3.96	696	L IFG (BA44)
COR-SEM	59	-52	32	3.68	471	R Angular gyrus
	50	-52	9	3.61	553	R Posterior MTG/STS

Talairach coordinates (*x*, *y*, *z*), Z-values, and volume (mm³) of the activated regions for the different contrasts: syntactically anomalous sentences vs. correct sentences (SYN-COR), semantically anomalous sentences vs. correct sentences (SEM-COR), and correct sentences vs. each anomalous condition (COR-SYN, COR-SEM). Z-values were thresholded at $Z > 3.09$ ($P < 0.001$, uncorrected) and clusters had a minimum size of 14 voxels (400 mm³). L1, native speakers of German; L2, non-native speakers of German; L, left; R, right; STG, superior temporal gyrus; IFG, inferior frontal gyrus.

more involved in the processing of correct versus syntactically anomalous sentences.

Semantically anomalous sentences, however, did bring on higher levels of activation than did correct sentences, specifically within the left IFG (BA44). This activation spread from superior regions of BA44 into inferior BA45/47. A direct contrast between correct versus semantically anomalous sentences showed increased levels of activation for correct sentences in right angular gyrus and right posterior superior temporal sulcus/middle temporal gyrus (STS/MTG).

Between-group comparison: non-native speakers versus native speakers. In between-group comparisons, the results show those areas that were activated differentially in each group (native speakers or non-native speakers of German) in response to the auditory presentation of well-formed syntactically anomalous and semantically anomalous German sentences. Talairach coordinates and the probability that the size of activation in a given area is reliably different between the groups based on Bayesian statistics can be found in Tables VI and VII. Time-courses depicting signal change over time and selected direct contrast maps can be seen in Figure 3. Mean signal changes in each ROI are depicted in Figure 4.

Direct contrasts

Non-native speakers versus native speakers. Non-native speakers showed a different pattern of activation than did native speakers in all three experimental conditions. When listening to well-formed, correct German sentences, non-native speakers showed a greater involvement of several

cortical and subcortical areas than that shown in native German speakers. Cortically, greater levels of increased activation were observed for non-native speakers in the left intraparietal sulcus, the left anterior insular cortex, and at three points within left frontal cortex. Frontal cortical activation was centered around three local maxima, located: (1) in the superior portion of BA44/6 at the junction point between inferior frontal sulcus and inferior precentral sulcus; (2) in a more inferior portion of BA44; and (3) in the left posterior orbital gyrus. On the subcortical level, non-native speakers showed a greater involvement of basal ganglia structures bilaterally, specifically in the head of the caudate nuclei.

When listening to syntactically incorrect sentences, non-native speakers showed a robust area of increased activation compared to that in native speakers in superior posterior reaches of the left IFG (BA44/6), and two smaller sites of cortical activation within the left intraparietal sulcus and right angular gyrus. Two further substantial sites of increased activation could be seen subcortically within the caudate nuclei bilaterally.

Semantically anomalous sentences brought on a small cortical activation in the left intraparietal sulcus in non-native speakers as compared to that in native speakers. Substantial activation could again be observed subcortically in the right and left caudate nuclei.

Native speakers versus non-native speakers. Native speakers of German listening to correct sentences in their native language showed greater levels of activation than did non-native speakers listening to the same sentences in the mid-

TABLE VI. Talairach coordinates, Z-values, volume and reliability of difference according to Bayes model of the activated regions for correct sentences in native speakers versus non-native speakers

Contrast	<i>x</i>	<i>y</i>	<i>z</i>	Z-max	Volume	Bayes (%)	Region
L2 > L1	-52	9	24	2.68	119	99.98	L IFG (BA 44/6)
	-49	6	6	3.01	95	99.99	L IFG (BA 44)
	-34	12	-12	2.97	196	99.95	L Posterior orbital gyrus
	-29	18	9	2.72	103	99.82	L Anterior insula
	-5	18	3	2.84	207	99.99	L Caudate nucleus
	11	15	6	3.22	508	99.99	R Caudate nucleus
	-28	-72	41	3.22	281	100	L Intraparietal sulcus
	-50	-21	12	3.59	868	99.99	L STG
	56	-39	12	2.88	206	100	R STG
	29	-27	0	2.96	149	99.89	R Temporal stem
L1 > L2	8	-66	35	3.45	722	100	R Precuneus
	25	21	12	2.91	512	99.98	R Anterior insula
	37	3	-6	3.20	396	99.74	R Anterior insula
	32	-33	18	3.38	1,446	100	R Posterior insula

Talairach coordinates (*x*, *y*, *z*), Z-values, volume (mm³) and reliability of difference according to Bayes model of the activated regions for non-native speakers (L2) vs. native speakers (L1) listening to correct sentences and L1 vs. L2 speakers listening to correct sentences. Z-values were thresholded at $Z > 2.32$ ($P < 0.01$, uncorrected).

BA, Brodmann's area; IFG, inferior frontal gyrus; STG, superior temporal gyrus.

portion of the bilateral STG, lateral to Heschl's gyrus, the right parietooccipital sulcus extending into the precuneus, and the right insular cortex.

Syntactically anomalous sentences brought on more activation in native than in non-native speakers in the mid-portion of the left STG and in several cortical sites within the

TABLE VII. Talairach coordinates, Z-values, volume and reliability of difference according to Bayes model of the activated regions for syntactically and semantically anomalous sentences in native and non-native speakers

Contrast	<i>x</i>	<i>y</i>	<i>z</i>	Z-max	Volume	Bayes (%)	Region
SYN							
L2 > L1	-52	9	24	3.16	1,104	100	L IFG
	-28	-72	41	3.34	499	100	L Intraparietal sulcus
	34	-57	35	3.44	308	100	R Angular gyrus (deep)
	-7	6	3	3.84	2,090	100	L Caudate nucleus
	8	12	9	3.29	875	99.99	R Caudate nucleus
L1 > L2	-53	-18	12	2.75	714	100	L Mid STG
	49	-12	12	2.48	232	99.99	R Mid STG
	55	-48	12	2.59	1,560	100	R Posterior STG
	32	-33	21	2.49	405	100	R Posterior insula
	22	-57	24	2.73	1,504	99.99	R Precuneus
	7	-54	15	2.62	299	99.99	R Posterior cingulate gyrus
SEM							
L2 > L1	-20	-78	32	2.91	262	99.99	L Intraparietal sulcus
	-5	6	3	3.47	1,760	99.99	L Caudate nucleus
	11	15	6	3.90	1,218	99.99	R Caudate nucleus
L1 > L2	-55	-18	9	3.49	762	99.99	L Mid STG
	52	-12	12	3.08	572	99.99	R Mid STG
	58	-39	12	2.98	1,120	99.99	R Posterior STS
	29	21	6	3.40	485	100	R Anterior insula
	35	-33	21	2.77	263	100	R Posterior insula
	8	-66	35	3.41	489	100	R Precuneus

Talairach coordinates (*x*, *y*, *z*), Z-values, volume (mm³) and reliability of difference according to Bayes model of the activated regions for the contrasts: non-native speakers (L2) vs. native speakers (L1) listening to syntactically (SYN) and semantically (SEM) anomalous sentences; L1 vs. L2 speakers listening to syntactically and semantically anomalous sentences. Z-values were thresholded at $Z > 2.32$ ($P < 0.01$, uncorrected).

IFG, inferior frontal gyrus; STG, superior temporal gyrus.

Figure 3.

Direct contrast maps of native vs. non-native speakers of German listening to correct (COR), semantically anomalous (SEM), and syntactically anomalous (SYN) sentences. Left inferior frontal gyrus (IFG; **A**) shows increased activation for non-native speakers of German in all conditions (see time-courses). Due to increased activation of IFG in the SEM condition in native speakers as well,

no significant difference is observed in IFG in the direct contrast between native and non-native speakers in the final panel. Bilateral caudate nucleus (**B**) is also activated significantly more by non-native speakers in all conditions. Left superior temporal gyrus (STG; **C**) shows greater levels of activation for native speakers than for non-native speakers.

right hemisphere. Right temporal lobe activation was observed in mid-portions of STG, homologous to the activation seen on the left. Furthermore, right posterior STG was shown to be more active in native speakers than in non-

native speakers. The right posterior insular cortex, right precuneus, and right posterior cingulate gyrus also showed increased levels of activation for native speakers compared to that for non-native speakers.

Figure 4.

Experiment 2: Mean percent signal change and standard error for native (L1) and non-native (L2) speakers of German in each of the ROIs discussed: inferior frontal gyrus (IFG); left caudate nucleus (LCN); right caudate nucleus (RCN); left superior temporal gyrus (LSTG); and right superior temporal gyrus (RSTG).

Semantic anomalies brought on increased levels of activation for native speakers in the STG bilaterally. In the left hemisphere, this activation was restricted to the mid portions of STG; in the right hemisphere, mid- and posterior portions of STG/STS were observed to show different levels of activation. In addition, the right anterior and posterior insular cortices, as well as the right parietooccipital sulcus spreading into precuneus regions, showed more activation in native speakers than in non-natives.

ROI analysis

A subsequent ROI analysis was conducted over five critical areas (left IFG [-49, 12, 6], left caudate nucleus [LCN; -5, 6, 3], right caudate nucleus [RCN; 11, 15, 6], left STG [LSTG; -50, -21, 12], and right STG [RSTG; 56, -39, 12]) to validate the statistical significance of observed activations. ANOVAs were calculated with the between-subjects factor group (L1 and L2), and within-subjects factors ROI (IFG, LCN, RCN, LSTG, and RSTG) and condition (COR, SEM, and SYN). Percent signal change was greater in both temporal ROIs than in the frontal or subcortical areas in both groups, resulting in a main effect of ROI ($F[4,120] = 25.26$; $P < 0.0001$). In each ROI and group, the condition COR showed the least activation, resulting in a main effect of condition ($F[2,60] = 4.23$; $P = 0.01$). The BOLD response elicited by the different conditions varied in the five ROIs as a factor of group, as reflected in the three-way interaction ROI $\times$ group $\times$ condition ($F[8,240] = 4.51$; $P < 0.001$).

The allowed post-hoc analysis of activation within each of the five ROIs provided the following results. Within the left IFG, the response of L2 speakers was greater than that of L1 speakers in all conditions, reflected in a main effect of group ($F[1,30] = 9.34$; $P < 0.005$). Furthermore, in both groups, the condition COR elicited the lowest rates of activation, as characterized by a main effect of condition ($F[2,60] = 8.39$; $P < 0.001$). Crucially, the two groups showed a different pattern of results across conditions, as reflected in the group $\times$ condition interaction ($F[2,60] = 6.29$; $P < 0.005$). This interaction can be explained by the fact that in the group L1, the condition SEM brought on significantly more activation than did either COR ($F[1,17] = 13.20$; $P < 0.005$) or SYN ($F[1,17] = 8.97$; $P < 0.01$), whereas in the group L2, the condition SYN brought on significantly more activation than did COR alone ($F[1,13] = 16.41$; $P = 0.001$). These values remained statistically significant after a Bonferroni significance level adjustment.

Within both subcortical ROIs (left and right caudate nucleus), L2 speakers showed a greater level of activation across all conditions, and no detectable differences between conditions. This was reflected in a main effect of group (LCN, $F[1,30] = 10.35$ and $P < 0.005$; RCN, $F[1,30] = 14.32$ and $P < 0.001$) with no further significant interactions.

Within the left STG, L1 speakers showed significantly more activation than did L2 speakers across conditions ($F[1,30] = 12.64$; $P < 0.005$). In both groups, the condition SYN brought on the highest level of activation ($F[2,60] =$

12.32; $P < 0.0001$). No further interaction reached significance.

Within the right STG, L1 speakers showed significantly more activation than did L2 speakers across conditions ($F[1,30] = 9.92$; $P < 0.005$). No main effect of condition and no group $\times$ condition interaction were observed.

DISCUSSION

Experiment I

We attempted to isolate specific language processing components (i.e., structural syntactic and semantic) from auditory language comprehension in general. To achieve this, participants listened to sentences that were correct, syntactically incorrect, or semantically incorrect. Changes in the hemodynamic response correlated with each sentence type were recorded and direct comparisons were calculated between the brain's response to each anomaly condition and its response to well-formed sentences. In this manner, brain regions activated selectively upon detection of syntactic or semantic errors could be identified. Such regions are not responsible for syntax or semantic processing per se, as both processes are clearly needed for the comprehension of correct sentences as well. All regions of increased activation common to an anomalous sentence and a correct sentence would not appear in a direct comparison of these two conditions. Rather, the regions seen activated in this study can be seen as investing extra resources upon identifying a problem in a given sentence. It is therefore important to point out that our experimental set-up was not designed to locate an exhaustive syntax or semantic network, but rather to identify those regions that become important upon detecting and processing specific types of linguistic errors, and if these errors are created in a similar manner in two different languages, whether or not similar processes are used by native speakers to detect said errors.

Superior temporal gyrus

In both groups of participants listening to syntactically anomalous versus correct sentences in their native language, increased activation was observed in left STG, centered around two neighboring foci: one located centrally, lateral to Heschl's gyrus, and one located more anteriorly within STG (see Fig. 1). That superior temporal cortex plays a role in language processing is relatively undisputed; however, its response to auditory stimuli other than language stimuli (i.e., tones) suggests that it is not an exclusively language-specific cortical area. In particular, those areas directly surrounding primary auditory cortex are suggested to support auditory processing in general (language included), whereas evaluation of highly complicated speech signals is dependent on recruitment of additional temporal areas (STS, MTG, and ITG) not needed for the perception of nonspeech cues [Binder et al., 2000]. For example, left STS/MTG in particular has been reported previously in semantic decision tasks [Binder et al., 1997], and the posterior reaches of the STS/

MTG have been postulated to reflect processes of sentence evaluation or sentential integration [Friederici et al., 2003]. The present results show no sites of increased activation in either group of native speakers in these areas. This is presumably due to the fact that analysis of correct sentences is equally dependent upon such processes, and that activation in such regions is therefore canceled out in a direct comparison.

The current results, however, do show increased activation correlating with syntactic phrase structure violations in lateral STG, anterior to Heschl's gyrus, on the supratemporal plane. This area has been cited in other studies examining online syntactic phrase structure building during auditory sentence comprehension [Friederici et al., 2000, 2003; Humphries et al., 2001; Meyer et al., 2000], and it has been suggested that a highly automatized local structure-building process is supported by this region. It is interesting that such a high degree of similarity exists between increased activation elicited by word category violations in two very different languages.

The second region of increased activation within temporal cortex, in the central portion of the left STG lateral to Heschl's gyrus, may be a reflection of processes not related directly to syntactic processing. The syntactically violated sentences were created in both German and Russian by inserting an incomplete prepositional phrase (PP) into an otherwise coherent sentence. Syntactically anomalous sentences were therefore always one word longer than correct sentences were (see examples in Table I). It is known that increased time spent on completing a given task brings on greater levels of activation in neuroimaging studies [Poldrack, 2000]. Along these same lines, it has been reported that an increasing amount of auditory input is correlated with increased activation in the STG [Binder and Price, 2001]. A second correct condition, which contained a completed PP, was presented to participants to ensure that error detection could not be based on the mere presence of a preposition. These sentences again are always longer than are the simple correct sentences, and allow us to test the hypothesis that the mid-STG activations observed in response to syntactically anomalous sentences are not necessarily a reflection of error detection. Indeed, correct sentences containing an additional PP also show increased activation in mid-STG bilaterally when directly contrasted with short correct sentences. Although we cannot say whether this increased activation is a reflection of the presentation of quantitatively more acoustic information (i.e., the extra PP) or the increased integration costs associated with incorporating this additional information into a simple sentence, we can say that mid-STG activation is not specific to the processing of syntactic violations.

Inferior frontal gyrus

The most robust site of increased activation for sentences containing a semantic violation in comparison to correct sentences could be seen in the anterior reaches of the IFG

(BA45/47; Fig. 1). Many studies looking at various aspects of semantic processing, specifically semantic retrieval, have reported left IFG activation [Cabeza and Nyberg, 2000; Dapretto and Bookheimer, 1999; Thompson-Schill et al., 1997; Wagner et al., 2001]. In particular, the inferior portion of IFG (BA47) has been suggested to play a role in processing semantic relationships between words or phrases, or in selecting a word based on semantic features from among competing alternatives [Bookheimer, 2002; Poldrack et al., 1999]. In the current study, participants faced with a semantically implausible word in a sentence experienced difficulties in establishing a sensible relationship between the anomalous word and the previous sentence context, resulting in increased levels of activation within inferior IFG. Importantly, such activation has nothing to do with long-term storage of semantic representations; rather, it is thought to reflect a very goal-oriented, strategic process of retrieval [Wagner et al., 2001] or comparison/analysis [Thompson-Schill et al., 1997]. It is only in the realization that a given word does not match the participant's expectations that such IFG activation makes sense in relation to semantic processes.

Experiment 2

The hemodynamic response of two different participant groups was recorded during auditory sentence presentation. The first of these groups was made up of highly proficient, late learners of German (native Russian speakers); the second group comprised native German speakers. We wished to investigate what differences, if any, could be observed in the cerebral activation of non-native versus native speakers listening to identical sentence materials, and to what extent different linguistic domains (syntactic processing vs. semantic processing) influence second-language processing.

We first address those areas shown to be more active in non-native speakers than in native speakers.

Non-native speakers

Frontal cortex. Non-native speakers listening to both correct and syntactically anomalous sentence stimuli showed several sites of increased activation in BA44 in comparison to that in native speakers. The local maximum of one of these activations was located within the superior posterior regions of BA44, along the anterior bank of the inferior precentral sulcus, and was observed in response to correct and syntactically anomalous sentences. The other was located inferior and anterior to this, also within BA44, and could be seen only in response to correct sentences.

The first of these regions corresponded to a portion of IFG cited in studies looking at strategic phonological processing [Burton et al., 2000; Poldrack et al., 1999]. Importantly, this area does not respond specifically to passive listening (i.e., does not support bottom-up stimulus-driven processes), but rather to strategic processing of auditory input. For example, phoneme discrimination tasks elicit increased activation in this area in comparison to that elicited by pitch discrimina-

tion tasks or passive listening to phonemes [Gandour et al., 2002; Zatorre et al., 1996]. Burton et al. [2000] argued that phoneme discrimination alone is not enough to elicit posterior inferior frontal gyrus (pIFG) activation; instead, tasks requiring segmentation of phonemes coupled with a discrimination task are needed to produce higher levels of activation. It is entirely plausible that non-native speakers experience increased difficulties in recognizing or categorizing acoustically presented phonemes within a speech signal. Behavioral studies seem to support this notion, as they have shown that age of acquisition of a second language influences phonological proficiency, in particular the perception of phonemes in noisy surroundings [Flege et al., 1999; Mayo et al., 1997; Meador et al., 2000]. In the current study, highly proficient but late learners of German were presented with acoustic sentence stimuli in the noisy scanner environment. The increased levels of activation for the non-natives observed in superior posterior IFG could well reflect the increased effort individuals in this group had to invest to perceive correctly and categorize the presented speech cues on a purely phonological level. The increased difficulty experienced by non-native speakers in all conditions is characterized by the behavioral results recorded: Non-native speakers made more errors than native speakers did in all experimental conditions.

Activation within IFG was different between native and non-native speakers listening to correct and syntactically anomalous sentences but not to semantically anomalous sentences. The absence of an observable difference between the groups for this condition is caused by the relative increase in IFG activation brought on by semantic anomalies in the group of native speakers. This is clear upon inspection of the time-course information and the ROI analysis. We attempt to account for this discrepancy: The perception of phonemes in spoken sentences presents no problem for native speakers, and correct sentence stimuli are understood easily. Non-native speakers, however, employ additional resources as outlined above to categorize phonemes correctly in even simple sentences. Syntactically anomalous sentences are disregarded quickly based on structural deficits and pose no further problem for native speakers. This is evident in the shorter reaction times for native speakers in response to syntactic errors and in the latency of ERP effects elicited by the same type of syntactic anomaly [ELAN, 150 ms; Hahne and Friederici, 2002]. There is also evidence that scanner noise does not affect the early syntactic processes as reflected in the ELAN [Herrmann et al., 2000]. Non-native speakers, however, experience increasing difficulties in this condition, and the activation under discussion is suggested to reflect difficulties in categorizing the phonemes of acoustically presented language materials, regardless of whether the stimuli is structurally correct or incorrect. In the case of semantic anomalies, however, even native speakers cannot rely on fast structural interpretations to determine the acceptability of a given sentence. In support of this assumption, semantic priming studies have shown that unrelated word pairs elicit greater activation in this area than do

related word pairs, which are integrated more easily [Kotz et al., 2002]. No indication of greater difficulty in detecting semantic anomalies can be detected in native speakers' error rates, although an indication of increased difficulty is perhaps reflected in reaction times, which showed significantly longer decision times for the detection of semantic anomalies than for syntactic anomalies or correct sentences. Although, as pointed out previously, reaction time measurements in this experiment should not be overinterpreted as they were not a reflection of true online sentence processing, it is an interesting observation to keep in mind.

A second area of interest within BA44 demonstrated greater levels of increased activation in non-native versus native speakers listening to correct sentence stimuli only. This portion of BA44 corresponds to previously reported findings concerning the processing of syntactic structure [Dapretto and Bookheimer, 1999; Fiebach et al., 2001; Friederici et al., 2000; Heim et al., 2003; Just et al., 1996]. In monolingual studies, an increased involvement of this region has been reported for the processing of sentences with increasing syntactic complexity [Caplan et al., 1998, 1999; Caplan, 2001; Just et al., 1996; Stromswold et al., 1996] and for the processing of syntactic transformations in particular [Ben-Shachar et al., 2003]. This brain area has also been implicated in the processing of syntactically incorrect sentences, but only when the error was set into focus by the task [Embick et al., 2000; Indefrey et al., 2001; Suzuki and Sakai, 2003]. In the current study, the fact that non-native participants showed more activation correlated with correct sentences than native speakers did in this region suggests that non-native speakers consistently engage more resources in syntactically parsing even simple, correct sentences in their second language. In other words, the lesser proficiency of non-native speakers in their second language causes even simple structures to be parsed as if they were complex. In the violation conditions, where native speakers also experience difficulties, no differences can be observed between native and non-native participants.

Caudate nucleus. Non-native speakers showed increased levels of activation in comparison to native speakers in the subcortical structures of the basal ganglia bilaterally for all sentence stimuli types. Specifically, increased levels of activation were observed in the head of the caudate nuclei in both hemispheres. Subcortical structures have been associated traditionally with coordination of movement, whereas studies looking at cognitive function have tended to concentrate on cortical activation. The indisputably crucial role of the basal ganglia in language processing, however, recently has begun to attract increasing attention [Lieberman, 2002; Stowe et al., 2003; Watkins et al., 2002b]. Clinical studies have shown that permanent loss of linguistic abilities associated with classic aphasia does not occur in the absence of subcortical damage [D'Esposito and Alexander, 1995; Dronkers et al., 1992; Lieberman, 2002]. Furthermore, focal damage to subcortical structures (for example in neurodegenerative illnesses such as Parkinson's disease) results in linguistic and cognitive deficits displaying properties of

classic aphasias [Lieberman, 2002]. In addition, developmental speech disorders have been shown to correlate with functional and structural abnormalities specifically in the caudate nucleus [Watkins et al., 2002b]. More recently, a number of ERP studies have shown that focal lesions of the basal ganglia show a selective deficit of controlled syntactic processes as reflected in the P600 [Friederici and Kotz, 2003; Frisch et al., 2003; Kotz et al., 2003].

The caudate nucleus works together with prefrontal cortex to support cognitive and linguistic function. Activation of the caudate nucleus should thus necessarily be correlated with prefrontal cortex activation. In the current study, non-native speakers showed increased levels of activation in comparison to native speakers in the head of the caudate nucleus bilaterally, but cortical activation only in left IFG. This is surprising, as functional neuroanatomy indicates that right caudate nucleus activity is tied necessarily to right hemispheric frontal cortical activity.

In light of this fact, we looked at the Z-maps for each individual participant and plotted local maxima within the left and right frontal cortices. The pattern of distribution seemed to be such that local maxima between subjects clustered around two areas within left IFG, corresponding roughly with those areas reported in the direct comparison of native and non-native speakers. In the right hemisphere, however, although most participants did show activation within prefrontal cortex (only 2 of 14 participants showed no local maximum in this region), the pattern of distribution was much more variable. In other words, whereas left hemispheric IFG activation seemed centered around two distinct foci, right hemispheric activation was present but quite dispersed. In evaluating group statistics, this can have a major effect. Although most participants showed activations in the same areas within left IFG, activations, although present, did not overlap in right IFG. This leads to a reliable effect in left hemisphere and no reliable effect in the right hemisphere. As the caudate nucleus is an anatomically much more restricted region, the chance that activation within caudate nucleus overlaps between participants is far greater than in frontal cortex. We postulate that non-native speakers thus show increased levels of activation in comparison to native speakers in left frontal language areas, and in homologous areas within the right hemisphere; however, right hemispheric activations are dispersed more widely and thus less robust in a group statistical analysis.

Native speakers

Superior temporal gyrus. Native speakers of German showed a greater involvement of temporal lobe areas than did non-native speakers when listening to all sentence types in German. Specifically, native speakers showed an area of increased activation in central portions of the STG, lateral to Heschl's gyrus (Fig. 3). STG and auditory association cortex clearly play a role in the processing of spoken speech, as in all auditory signals, although a functional breakdown of regions within STG in relation to language processing has not yet been determined clearly. It has been argued that

anterior and posterior regions of STG specifically process speech signals [Giraud and Price, 2001; Scott et al., 2000; Scott and Johnsrude, 2003]. Anterior portions of STG have been suggested to support very specific, semantically driven aspects of lexical retrieval [Binder and Price, 2001; Kiehl et al., 2002] and also syntactically motivated phrase-structure building processes [Friederici et al., 2000, 2003; Humphries et al., 2001; Meyer et al., 2000], whereas posterior STG has been suggested to support general sentential evaluation and integration [Friederici et al., 2003].

The area in which native speakers show significantly more involvement than do non-native speakers in processing spoken sentences lies within the mid-portion of STG, in neither of the two functionally delineated areas of temporal cortex. We suggest that the increased level of activation observed in this area in native speakers reflects highly automatized and efficient processes of acoustic phonological processing (as mentioned above for Experiment 1). Non-native speakers have difficulties efficiently decoding the incoming speech signal on the phonological level, and are thus forced to employ additional resources in IFG, a process not observed in native speakers. Native speakers, however, have no need to strategically analyze incoming phonemes belonging to their familiar native language.

General Discussion and Conclusions

The study introduced here had two aims. First, we wished to explore the neural correlates underlying syntactic and semantic processing across different native languages. We attempted to achieve this by comparing the hemodynamic response of a group of participants to both syntactically and semantically anomalous sentences in their native language and then subtracting from this the hemodynamic response correlated with correct sentence processing. In this manner, we hoped to see those areas specifically engaged in detecting problems in the domains of syntactic and semantic integration. Our results indicate that syntactic processing is supported by a frontotemporal network, with specific increased involvement of anterior portions of STG. This finding is in keeping with several other studies looking at online auditory sentence processing [Humphries et al., 2001; Meyer et al., 2000]. Semantic processing, however, selectively recruits portions of the IFG bilaterally, although with a clear dominance in the left hemisphere. Specifically, the anterior portion of the left IFG (BA45/47) seems involved in processing of semantic anomalies. Again, this finding is in keeping with current neuroimaging literature [Bookheimer, 2002; Dapretto and Bookheimer, 1999].

Comparable linguistic materials in two typologically different languages (German and Russian) were shown to elicit comparable brain responses in native speakers of each respective language. Because previous neuroimaging studies investigating language processing have been conducted in different languages [i.e., Italian, Perani et al., 1996; Chinese, Luke, 2002; Thai, Gandour et al., 2003; Spanish, Perani et al., 1998; Catalan, Rodriguez-Fornells et al., 2002; German, Friederici et al., 2003], and those areas observed to be in-

volved in various aspects of language processing seem to overlap, we expected to see no difference between German natives processing German sentence stimuli and Russian native speakers processing Russian stimuli. To our knowledge, there is no existing study directly comparing auditory sentence comprehension in two different native populations. Our results show comparable sites of increased activation for the two groups of native speakers.

The second goal of the present study was to investigate to what extent native- and non-native-language processing differ from one another. To this end, we presented non-native speakers of German (native speakers of Russian) with spoken German sentences. Crucially, we have shown already that native speakers of Russian listening to similarly constructed materials in their native language demonstrated similar patterns of activation to German natives. This allowed us to assume that the underlying language processing strategies employed to overcome the given types of linguistic violations were similar for Russian and German native speakers. Any difference observed in the processing of German by non-native speakers was thus a reflection of different processing strategies between L1 and L2 rather than different biologically determined language networks per se.

We observed different patterns of activation in non-native speakers of German presented with auditory sentence materials in German. Non-native speakers showed no reliable differences in the processing of syntactically anomalous versus correct sentences. We postulate that this was a reflection of the ever-present increased syntactic processing costs associated with parsing a second language. Non-native speakers did, however, show at least partially comparable activations to native speakers in response to semantically anomalous sentences versus correct sentences. Overlapping activation between the groups was observed in the left anterior IFG, a cortical area typically observed in semantic processing studies [Bookheimer, 2002; Poldrack et al., 1999; Wagner et al., 2001]. In summary, we can say that processes underlying semantic processing seemed to be more similar in native and non-native speakers of a given language than did those underlying syntactic processing.

In a final processing step, we directly compared the increased brain activation elicited in the group of native versus non-native participants in response to each experimental sentence type. In this manner we were able to show those brain areas showing a differential response to native versus non-native language processing in general. In all experimental conditions, non-native speakers showed greater activation than native speakers did within the head of the caudate nuclei bilaterally. The role of basal ganglia structures in language processing has begun to receive increasing attention [Lieberman, 2002], and functional and structural abnormalities of the caudate nucleus have been shown to be characteristic of developmentally language-impaired patients [Watkins et al., 2002a,b]. Furthermore, non-native speakers showed a reliable region of increased activation in comparison to native speakers within IFG. This activation

increase was observed in response to correct and syntactically incorrect sentences but not to semantically incorrect sentences. Based on previous studies showing a correlation between posterior IFG activation and phonological processing [Bookheimer, 2002; Poldrack et al., 1999], we postulate that the observed activation increase for non-native speakers reflected additional phonological processing necessary for understanding speech signals in a foreign language. L2 speakers seem to compensate for less proficient phonological decoding at the level of STG by additionally recruiting frontal cortex. In the case of semantic anomalies, we suggest that native speakers also attempt to find a phonologically similar, semantically appropriate alternative to the presented violation, and thus demonstrate no difference to non-native speakers.

ACKNOWLEDGMENTS

We thank A. Hahne for providing the German sentence materials, S. Zysset, K. Müller, and J. Neumann for invaluable support with the data analysis, C. Preul and D.Y. von Cramon for discussions, and M. Naumann, A. Wiedemann, and S. Wipper for help with data acquisition.

REFERENCES

- Abutalebi J, Cappa S, Perani D (2001): The bilingual brain as revealed by functional neuroimaging. *Bilingualism: Language and Cognition* 4:179–190.
- Angrilli A, Penolazzi B, Vespignani F, De Vincenzi M, Job R, Ciccarelli L, Palomba D, Stegagno L (2002): Cortical brain responses to semantic incongruity and syntactic violation in Italian language: an event-related potentials study. *Neurosci Lett* 322:5–8.
- Ben-Shachar M, Hendler T, Kahn I, Ben-Bashar D, Grodzinsky Y (2003): The neural reality of syntactic transformations: evidence from functional magnetic resonance imaging. *Psychol Sci* 14: 433–440.
- Binder J, Frost J, Hammeke T, Cox R, Rao S, Prieto T (1997): Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci* 17:353–362.
- Binder J, Frost J, Hammeke T, Bellgowan P, Springer J, Kaufman J, Possing E (2000): Human temporal lobe activation by speech and nonspeech sounds. *Cereb Cortex* 10:512–528.
- Binder J, Price C (2001): Functional neuroimaging of language. In: Cabeza R, Kingstone A, editors. *Handbook of functional neuroimaging of cognition*. Cambridge, MA: Bradford Books. p 187–252.
- Bookheimer S (2002): Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annu Rev Neurosci* 25:151–188.
- Burton M, Small S, Blumstein S (2000): The role of segmentation in phonological processing: an fMRI investigation. *J Cogn Neurosci* 12:679–690.
- Cabeza R, Nyberg L (2000): Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J Cogn Neurosci* 12:1–47.
- Caplan D (2001): Functional neuroimaging studies of syntactic processing. *J Psycholinguist Res* 30:297–320.
- Caplan D, Alpert N, Waters G (1998): Effects of syntactic structure and propositional number on patterns of regional blood flow. *J Cogn Neurosci* 10:541–552.

- Caplan D, Alpert N, Waters G (1999): PET studies of sentence processing with auditory sentence presentation. *Neuroimage* 9:343–351.
- Chee M, Caplan D, Soon C, Sriram N, Tan E, Thiel T, Weekes B (1999a): Processing of visually presented sentences in Mandarin and English studied with fMRI. *Neuron* 23:127–137.
- Chee M, Tan E, Thiel T (1999b): Mandarin and English single word processing studied with functional magnetic resonance imaging. *J Neurosci* 19:3050–3056.
- Chee M, Weekes B, Lee K, Soon C, Schreiber A, Hoon J, Chee M (2000): Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI. *Neuroimage* 12:392–403.
- Connolly J, Phillips N (1994): Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. *J Cogn Neurosci* 6:256–266.
- Cooke A, Zurif E, DeVita C, Alsop D, Koenig P, Detre J, Gee J, Pingo M, Balogh J, Grossman M (2001): Neural basis for sentence comprehension: grammatical and short-term memory components. *Hum Brain Mapp* 15:80–94.
- Dapretto M, Bookheimer S (1999): Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron* 24:427–432.
- Dehaene S, Dupoux E, Mehler J, Cohen L, Paulesu E, Perani D, van de Moortele P, Lehericy S, LeBihan D (1997): Anatomical variability in the cortical representation of first and second language. *Neuroreport* 8:3809–3815.
- D'Esposito M, Alexander M (1995): Subcortical aphasia: distinct profiles following left putaminal hemorrhage. *Neurology* 45:38–41.
- Deutsch A, Bentin S (2001): Syntactic and semantic factors in processing gender agreement in Hebrew: evidence from ERPs and eye movement. *J Mem Lang* 45:200–224.
- Dronkers N, Shapiro N, Redfern B, Knight R (1992): The role of Broca's area in Broca's aphasia. *J Clin Exp Neuropsychol* 14:198.
- Embick D, Marantz A, Miyashita Y, O'Neil W, Sakai K (2000): A syntactic specialization for Broca's area. *Proc Natl Acad Sci USA* 97:6150–6154.
- Fiebach CJ, Schlesewsky M, Friederici AD (2001): Syntactic working memory and the establishment of filler-gap dependencies: insights from ERPs and fMRI. *J Psycholinguist Res* 30:321–338.
- Fiebach CJ, Schlesewsky M, Lohmann G, Von Cramon DY, Friederici AD (2004): Revisiting the role of Broca's area in sentence processing: syntactic integration vs. syntactic working memory. *Hum Brain Mapp* 24:79–91.
- Flege J, MacKay I, Meador D (1999): Native Italian speakers' perception and production of English vowels. *J Acoust Soc Am* 106:2973–2987.
- Fox P, Ingham R, Ingham J, Hirsch T, Downs JH, Martin C, Jerabek P, Glass T, Lancaster J (1996): A PET Study of the neural systems of stuttering. *Nature* 382:158–162.
- Friederici AD (2002): Towards a neural basis of auditory sentence processing. *Trends Cogn Sci* 6:78–84.
- Friederici AD, Hahne A, Mecklinger A (1996): The temporal structure of syntactic parsing: early and late event-related brain potential effects elicited by syntactic anomalies. *J Exp Psychol Learn Mem Cogn* 22:1219–1248.
- Friederici AD, Kotz S (2003): The brain basis of syntactic processes: functional imaging and lesion studies. *Neuroimage* 20(Suppl.): 8–17.
- Friederici AD, Meyer M, von Cramon DY (2000): Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. *Brain Lang* 74:289–300.
- Friederici AD, Rüschemeyer S-A, Hahne A, Fiebach CJ (2003): The role of left inferior frontal and superior temporal cortex in sentence comprehension. *Cereb Cortex* 13:170–177.
- Frisch S, Kotz S, von Cramon D, Friederici AD (2003): Why the P600 is not just a P300: the role of the basal ganglia. *Clin Neurophysiol* 114:336–340.
- Friston K, Fletcher P, Josephs O, Holmes A, Rugg M, Turner R (1998): Event-related fMRI: characterizing differential responses. *Neuroimage* 7:30–40.
- Gandour J, Wong D, Lowe M, Dziedzic M, Saththamnuwong N, Tong Y, Xiaojian L (2002): A cross-linguistic fMRI study of spectral and temporal cues underlying phonological processing. *J Cogn Neurosci* 14:1076–1087.
- Giraud A, Price C (2001): The constraints functional neuroimaging places on classical models of auditory word processing. *J Cogn Neurosci* 13:754–765.
- Hagoort P, Brown CM (2000): ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia* 38:1531–1549.
- Hahne A (2001): What's different in second-language processing? Evidence from event-related brain potentials. *J Psycholinguist Res* 30:251–265.
- Hahne A, Friederici AD (1999): Electrophysiological evidence for two steps in syntactic analysis: early automatic and late controlled processes. *J Cogn Neurosci* 11:194–205.
- Hahne A, Friederici AD (2002): Differential task effects on semantic and syntactic processes as revealed by ERPs. *Brain Res Cogn Brain Res* 13:339–356.
- Heim S, Opitz B, Friederici AD (2003): Distributed cortical networks for syntax processing: Broca's area as the common denominator. *Brain Lang* 85:402–408.
- Herrmann C, Oertel U, Wang Y, Maess B, Friederici AD (2000): Noise affects auditory and linguistic processing differently: an MEG study. *Neuroreport* 1:227–229.
- Humphries C, Kimberley T, Buchsbaum B, Hickok G (2001): Role of anterior temporal cortex in auditory sentence comprehension: an fMRI study. *Neuroreport* 12:1749–1752.
- Indefrey P, Hagoort P, Herzog H, Seitz R, Brown C (2001): Syntactic processing in left prefrontal cortex is independent of lexical meaning. *Neuroimage* 14:546–555.
- Josephs O, Turner R, Friston K (1997): Event-related fMRI. *Hum Brain Mapp* 5:243–248.
- Just M, Carpenter P, Keller T, Eddy W, Thulborn K (1996): Brain activation modulated by sentence comprehension. *Science* 274: 114–116.
- Kaan E, Harris A, Gibson E, Holcomb P (2000): The P600 as an index of syntactic integration difficulty. *Lang Cogn Process* 15:159–201.
- Kiehl A, Laurens K, Liddle P (2002): Reading anomalous sentences: an event-related fMRI study of semantic processing. *Neuroimage* 17:842–850.
- Kim K, Relkin N, Lee K, Hirsch J (1997): Distinct cortical areas associated with native and second languages. *Nature* 388:171–174.
- Kotz S, Cappa S, von Cramon DY, Friederici AD (2002): Modulation of the lexical-semantic network by auditory semantic priming: an event-related functional MRI study. *Neuroimage* 17:1761–1772.

- Kotz S, Frisch S, von Cramon DY, Friederici AD (2003): Syntactic language processing: ERP lesion data on the role of the base ganglia. *J Int Neuropsychol Soc* 9:1053–1060.
- Kuperberg G, McGuire P, Bullmore E, Brammer M, Rabe-Hesketh S, Wright I, Lythgoe D, Williams S, David A (2000): Common and distinct neural substrates for pragmatic, semantic and syntactic processing of spoken sentences: an fMRI study. *J Cogn Neurosci* 12:321–341.
- Kutas M, Van Petten C (1994): Psycholinguistics electrified. Event-related brain potential investigations. In: Gernsbacher MA, editor. *Handbook of psycholinguistics*. San Diego: Academic Press. p 83–143.
- Lee J, Garwood M, Menon R, Adriany G, Andersen P, Truwit C, Ugurbil K (1995): High contrast and fast three dimensional magnetic resonance imaging at high fields. *Magn Reson Med* 34:308.
- Lieberman P (2002): On the nature and evolution of the neural bases of human language. *Yearb Phys Anthropol* 45:36–62.
- Lohmann G, Mueller K, Bosch V, Mentzel H, Hessler S, Chen L, Zysset S, von Cramon DY (2001): Lipsia—a new software system for the evaluation of functional magnetic resonance images of the human brain. *Comput Med Imaging Graph* 25:449–457.
- Luke K, Liu H, Wai Y, Wan Y, Tan L (2002): Functional anatomy of syntactic and semantic processing in language comprehension. *Hum Brain Mapp* 16:133–145.
- Mayo L, Florentine M, Buus S (1997): Age of second-language acquisition and perception of speech in noise. *J Speech Lang Hear Res* 40:686–693.
- Meador D, Flege J, MacKay I (2000): Factors affecting the recognition of words in a second language. *Bilingualism Lang Cogn* 3:55–67.
- Meyer M, Friederici AD, von Cramon DY (2000): Neurocognition of auditory sentence comprehension: event-related fMRI reveals sensitivity to syntactic violations and task demands. *Brain Res Cogn Brain Res* 9:19–33.
- Miezin F, Maccotta L, Ollinger J, Petersen S, Buckner R (2000): Characterizing the hemodynamic response: effects of presentation rate, sampling procedure and the possibility of ordering brain activity based on relative timing. *Neuroimage* 11: 735–759.
- Moro A, Tettamanti M, Perani D, Donati C, Cappa S, Fazio F (2001): Syntax and the brain: disentangling grammar by selective anomalies. *Neuroimage* 13:110–118.
- Mueller JL (in press): Electrophysiological correlates of second language processing. *Second Language Research*.
- Näätänen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, Vainio V, Alku P, Ilmoniemi R, Luuk A, Allik J, Sinkkonen J, Alho K (1997): Language-specific phoneme representations revealed by electric and magnetic brain responses. *Nature* 385:432–434.
- Nakada T, Fujii Y, Kwee I (2001): Brain strategies for reading in the second language are determined by the first language. *Neurosci Res* 40:351–358.
- Nakagome K, Takazawa S, Kanno O, Hagiwara H, Nakajima H, Itoh K, Koshida I (2001): A topographical study of ERP correlates of semantic and syntactic violations in the Japanese language using the multichannel EEG system. *Psychophysiology* 38:304–315.
- Neumann J, Lohmann G (2003): Bayesian second-level analysis of functional magnetic resonance images. *Neuroimage* 20:1346–1355.
- Ni W, Constable R, Mencl W, Pugh K, Fulbright R, Shaywitz S, Shaywitz B, Gore J, Shankweiler D (2000): An event-related neuroimaging study distinguishing form and content in sentence processing. *J Cogn Neurosci* 12:120–133.
- Norris D (2000): Reduced power multi-slice MDEFT imaging. *J Magn Reson Imaging* 11:445–451.
- Oldfield R (1971): The assessment and analysis of handedness: the Edinburgh Inventory. *Neuropsychologia* 9:97–113.
- Osterhout L, Holcomb P, Swinney D (1994): Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. *J Exp Psychol Learn Mem Cogn* 20:786–803.
- Pallier C, Dehaene S, Poline JB, LeBihan D, Argenti AM, Dupoux E, Mehler J (2003): Brain imaging of language plasticity in adopted adults: can a second language replace the first? *Cereb Cortex* 13:155–161.
- Paulesu E, McCrory E, Menoncello L, Brunswick N, Cappa S, Cotelli M, Cossu G, Corte F, Lorusso M, Pesenti S, Gallagher A, Perani D, Price C, Frith C, Frith U (2000): A cultural effect on brain function. *Nat Neurosci* 3:91–96.
- Perani D, Abutalebi J, Paulesu E, Brambati S, Scifo P, Cappa S, Fazio F (2003): the role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Hum Brain Mapp* 19:170–182.
- Perani D, Dehaene S, Grassi F, Cohen L, Cappa S, Dupoux E, Fazio F, Mehler J (1996): Brain processing of native and foreign languages. *Neuroreport* 7:2439–2444.
- Perani D, Paulesu E, Galles N, Dupoux E, Dehaene S, Bettinardi V, Cappa S, Fazio F, Mehler J (1998): The bilingual brain: proficiency and age of acquisition of the second language. *Brain* 121:1841–1852.
- Poldrack R (2000): Imaging brain plasticity: conceptual and methodological issues—a theoretical review. *Neuroimage* 12:1–13.
- Poldrack R, Wagner A, Prull M, Desmond J, Glover G, Gabrieli J (1999): Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage* 10:15–35.
- Price C (2000): The anatomy of language: contributions from functional neuroimaging. *J Anat* 197:335–359.
- Rodriguez-Fornells A, Rotte M, Heinze H-J, Nösselt T, Münte T (2002): Brain potential and functional MRI evidence for how to handle two languages with one brain. *Nature* 415:1026–1029.
- Scott S, Blank C, Rosen S, Wise R (2000): Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123:2400–2406.
- Scott S, Johnsrude I (2003): The neuroanatomical and functional organization of speech perception. *Trends Neurosci* 26:100–107.
- Stowe L, Paans A, Wijers A, Zwarts F (2003): Activations of “motor” and other non-language structures during sentence comprehension. *Brain Lang* 89:290–299.
- Stromswold K, Caplan D, Alpert N, Rauch S (1996): Localization of syntactic comprehension by positron emission tomography. *Brain Lang* 52:452–473.
- Suzuki K, Sakai K (2003): An event-related fMRI study of explicit syntactic processing of normal/anomalous sentences in contrast to implicit syntactic processing. *Cereb Cortex* 13:517–526.
- Talairach P, Tournoux J (1988): A stereotactic coplanar atlas of the human brain. Stuttgart: Thieme.
- Tan L, Spinks J, Feng C, Siok W, Perfetti C, Xiong J, Fox P, Gao J (2003): Neural systems of second language reading are shaped by native language. *Hum Brain Mapp* 18:158–166.
- Thirion J (1998): Image matching as a diffusion process: an analogy with Maxwell’s demons. *Med Image Anal* 3:243–260.

- Thompson-Schill S, D'Esposito M, Aguirre G, Farah M (1997): Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proc Natl Acad Sci USA* 94:14792–14797.
- Wagner A, Pare-Blagoev E, Clark J, Poldrack R (2001): Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron* 31:329–338.
- Wartenburger I, Heekeren H, Abutalebi J, Cappa S, Villringer A, Perani D (2003): Early setting of grammatical processing in the bilingual brain. *Neuron* 37:159–170.
- Watkins K, Dronkers N, Vargha-Khadem F (2002a): Behavioural analysis of an inherited speech and language disorder: comparison with acquired aphasia. *Brain* 125:451–463.
- Watkins K, Vargha-Khadem F, Ashburner J, Passingham RE, Connelly A, Friston KJ, Frackowiak RSJ, Mishkin V, Gadian DG (2002b): MRI analysis of an inherited speech and language disorder: structural brain abnormalities. *Brain* 125:465–478.
- Worsley KJ, Friston KJ (1995): Analysis of fMRI time-series revisited—again. *Neuroimage* 2:173–181.
- Yetkin O, Yetkin Z, Haughton V, Cox R (1996): Use of functional MR to map language in multilingual volunteers. *Am J Neuroradiol* 17:473–477.
- Zatorre RJ, Meyer E, Gjedde A, Evans AC (1996): PET Studies of phonetic processing of speech: review, replication and reanalysis. *Cereb Cortex* 6:21–30.