# Neural Correlates of the Perception of Contrastive Prosodic Focus in French: A Functional Magnetic Resonance Imaging Study

Marcela Perrone-Bertolotti,[1] Marion Dohen,[2] Hélène Lœvenbruck,[2] Marc Sato,[2] Cédric Pichat,[1] and Monica Baciu[1]*

[1]Laboratoire de Psychologie et NeuroCognition, UMR CNRS 5105, Université Pierre Mendès-France, Grenoble, France
[2]Département Parole et Cognition, GIPSA-Lab, UMR CNRS 5216, Grenoble Université, Grenoble, France

◆━━━━━━━━━━━━━━━━━━━━━━━━━━━◆
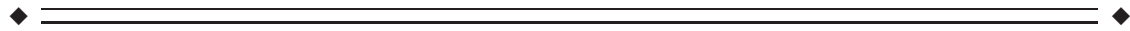
Abstract: This functional magnetic resonance imaging (fMRI) study aimed at examining the cerebral regions involved in the auditory perception of prosodic focus using a natural focus detection task. Two conditions testing the processing of simple utterances in French were explored, narrow-focused versus broad-focused. Participants performed a correction detection task. The utterances in both conditions had exactly the same segmental, lexical, and syntactic contents, and only differed in their prosodic realization. The comparison between the two conditions therefore allowed us to examine processes strictly associated with prosodic focus processing. To assess the specific effect of pitch on hemispheric specialization, a parametric analysis was conducted using a parameter reflecting pitch variations specifically related to focus. The comparison between the two conditions reveals that brain regions recruited during the detection of contrastive prosodic focus can be described as a right-hemisphere dominant dual network consisting of (a) ventral regions which include the right posterosuperior temporal and bilateral middle temporal gyri and (b) dorsal regions including the bilateral inferior frontal, inferior parietal and left superior parietal gyri. Our results argue for a dual stream model of focus perception compatible with the asymmetric sampling in time hypothesis. They suggest that the detection of prosodic focus involves an interplay between the right and left hemispheres, in which the computation of slowly changing prosodic cues in the right hemisphere dynamically feeds an internal model concurrently used by the left hemisphere, which carries out computations over shorter temporal windows. *Hum Brain Mapp* 34:2574–2591, 2013. © 2012 Wiley Periodicals, Inc.

Key words: prosody; contrastive focus; hemispheric specialization; speech perception; ventral and dorsal networks; internal models

◆━━━━━━━━━━━━━━━━━━━━━━━━━━━◆
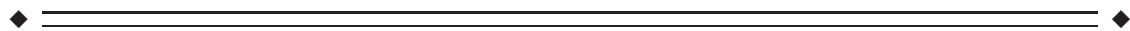
INTRODUCTION

Prosody (intonation, rhythm, and phrasing) is crucial in language processing. In everyday conversation, prosody plays a critical role in signaling both paralinguistic (e.g., attitudes such as sympathy, politeness, irony, or sarcasm) and affective (i.e., the emotional state of the speaker, e.g., anger, fear, happiness, surprise, sadness, disgust) aspects of communication as well as various different linguistic ones such as segmentation (into words or phrases),

utterance type (statement, question, or command), lexical stress, focus, and type of speech act. The neural correlates of its perception still remain unclear, however, due to its multiple functions and modes of expression as well as its multiple interactions with segmental, lexical, syntactic [Selkirk, 1978], and semantic [Ladd, 1996] processing in sentence comprehension.

In the past decades, research has largely focused on the respective roles of the left (LH) and right (RH) cerebral hemispheres in the processing of prosodic information. Based on dichotic listening, lesion deficits, and brain imaging data, an RH lateralization hypothesis was proposed [e.g., lesion data: Brådvik et al., 1991; Twist et al., 1991; Weintraub et al., 1981; dichotic listening: Shipley-Brown et al., 1988; brain imaging: Hesling et al., 2005; Meyer et al., 2003; Zatorre et al.,1992]. However, this RH lateralization hypothesis was challenged by a number of brain imaging studies, especially when linguistic, rather than paralinguistic and affective, aspects were taken into account [Astesano et al., 2004; Stiller et al., 1997; Tong et al., 2005]. It therefore appears that the processing of prosody cannot be restricted to one hemisphere. Several factors may influence hemispheric asymmetry and further hypotheses have been formulated to explain the role played by each hemisphere in the processing of prosody.

The first hypothesis, called functional lateralization hypothesis, suggests a purely functional hemispheric specialization: the perception of affective (or emotional) prosody would be primarily mediated by the RH while linguistic prosody would be bound to the LH [Gandour et al., 2003b; Luks et al., 1998; Pell, 1999; Wildgruber et al., 2004]. A recent functional magnetic resonance imaging (fMRI) study [Wiethoff et al., 2008], however, suggests that the RH dominance for the processing of affective prosody would rather be due to "an interplay of acoustic cues which express emotional arousal in the speaker's voice" than to a functional distinction between affective and linguistic prosody.

A second hypothesis, which we will refer to as the acoustic cue lateralization hypothesis, suggested that hemispheres would rather be specialized in the processing of specific cues. Several assumptions were made. A first assumption suggests that the RH is predominant for the processing of acoustic prosodic cues in general while the LH is predominant for the linguistic processing of these cues [Gandour et al., 2004; Perkins et al., 1996; Wildgruber et al., 2006]. A second assumption proposes a differential processing of the acoustic cues: fundamental frequency (F_0) would depend on the RH while duration would be processed in the LH [Van Lancker and Sidtis, 1992; Zatorre et al., 2002]. The results are still controversial, however, [Baum, 1998; Pell, 1998; Pell and Baum, 1997; Schirmer et al., 2001] a number of studies have challenged these assumptions. Differential hemisphere involvement has indeed been found for the same auditory signal (i.e., same sound, same acoustic cues), according to the functional context and/or linguistic relevance. More specifically,

Geiser et al. [2008] have shown task-related asymmetry in the processing of metrical patterns. The same stimuli were processed in the RH during an explicit rhythm task but were handled by the LH during a prosody task. Similarly, Meyer et al. [2005] have shown different hemispheric involvement depending on processing mode, using a set of sine-wave analogues that could be perceived as either non-speech or speech. They only observed an activation increase in the left posterior primary and secondary auditory cortex, when the stimuli were perceived as speech. In the same vein, Shtyrov et al. [2005] examined magnetic brain responses to the same stimulus, placed in contexts where it was perceived either as a non-speech noise burst, as a phoneme in a pseudoword or as a phoneme in the context of words. Left hemispheric dominance was only found when the sound was placed in word context. Furthermore, hemispheric asymmetry has been observed for the processing of the same speech signal depending on the language experience of the listener [standard Dutch vs. tonal dialect of Roermond: Fournier et al., 2010; Chinese vs. English: Gandour et al., 2003a; Chinese vs. Thai: Xu et al., 2006].

Finally, another hypothesis, referred to as the Asymmetric Sampling in Time (AST), is that hemispheric specialization would depend on the timescale over which the acoustic prosodic cues are analyzed. A speech signal consists of rapidly changing (formant transitions, bursts, below the phoneme domain) and slowly changing (fundamental frequency changes over the syllable, the phrase or the utterance) acoustic features. A rightward asymmetry has been found in the auditory cortex for slow relative to rapid formant transitions [e.g. Belin et al., 1998; Boemio et al., 2005; Giraud et al., 2007; Overath et al., 2008]. The RH would be crucial for analyses over longer timescales (150–300 ms), whereas analyses over shorter timescales (20–50 ms) would be done bilaterally [e.g., Hickok and Poeppel, 2007; Poeppel, 2003]. This hypothesis, is partly supported by a number of studies on prosody, showing that more local cues (such as tones) are processed in the LH and global intonation contours in the RH [Fournier et al., 2010; Gandour et al., 2003a; Meyer et al., 2002, 2004].

In sum, what all these studies put forward is that linguistic prosody is processed by a complex neural network of brain regions involving both hemispheres. The two hemispheres could be involved differentially depending on functional requirements (e.g., rhythm task vs. speech task) and/or on the nature of the acoustic cues (duration vs. fundamental frequency cues or linguistic vs. non-speech specific cues), on language experience (native vs. non native) and on temporal integration windows (long vs. short timescales). The fact that prosodic cues can spread over different domains (syllable vs. word vs. entire utterance) is also a problem because it appears that auditory areas in both hemispheres could be specialized in the processing of cues extending over short vs. long timescales.

This study aims at assessing the neural correlates of the perception of a specific linguistic prosodic phenomenon,

namely contrastive focus, and their potential hemispheric specialization. Focus in linguistics refers to that part of a sentence, which expresses the centre of attention of the utterance, that part of its meaning which is not presupposed in discourse. The focus constituent may, in Ladd's [1980] terminology, be "broad" or "narrow," depending on size. Contrastive narrow focus highlights a constituent within an utterance without change to the segmental content, thereby signaling to the listener that a given constituent is the most informative in the utterance (e.g., "Did Carol eat the apple? No, SARAH ate the apple," where capital letters signal contrastive focus). Contrastive narrow focus therefore makes an exclusive selection of a unit inside a paradigmatic class. It is of particular interest in language processing because it conveys an important pragmatic function (namely, referential identification or pointing) and is frequently used in natural conversation. This type of narrow focus is referred to variously as contrastive focus, identificational focus, alternatives focus, corrective focus, or simply focus [for reviews, see Gussenhoven, 2007; Selkirk, 2008]. For sake of simplicity, it will be referred to hereafter as "focus." Although it is well known that the focused constituent bears a number of recognizable intonational and durational cues affecting the entire utterance [for French: e.g., Di Cristo, 2000; Dohen and Lœvenbruck, 2004], previous studies on the articulatory [Dohen et al., 2004, 2009] and neural correlates of prosodic focus production [Lœvenbruck et al., 2005] also suggest that it involves accurate sensorimotor representations of how the articulators should be positioned. Importantly, these sensorimotor representations do not seem to be involved when focus is achieved through other means such as syntax, using cleft constructions [Lœvenbruck et al., 2005].

Regarding the neural correlates of the processing of prosodic focus, only a few studies have been carried out and they have led to mixed results. Wildgruber et al. [2004] compared the perception of prosodic focus (linguistic prosody condition) with that of affective (emotional) prosody using fMRI. Sentences varying in focus location (e.g. "the SCARF is in the chest" vs. "the scarf is in the CHEST") as well as emotional expressiveness (e.g., simple statement vs. excited statement) were generated by systematic manipulations of the F_0 contour of a simple declarative sentence. A discrimination task on pairs of F_0 resynthesized stimuli was used. The task either involved discrimination of focus location (linguistic prosody) or discrimination of emotional expressiveness (affective prosody). As compared with a resting condition, both conditions yielded bilateral hemodynamic responses within supplementary motor area, anterior cingulate gyrus, superior temporal gyrus, frontal operculum, anterior insula, thalamus, and cerebellum. Responses within the dorsolateral frontal cortex (BA 9/45/46) showed right lateralization effects during both tasks. The authors concluded that extraction and comparison of pitch patterns are mediated by the dorsolateral prefrontal cortex, the frontal opercu-

lum, the anterior insula and the superior temporal cortex and appear partially right-lateralized. Linguistic and emotional prosody were also directly compared with each other. The authors found that perception of prosodic focus (vs. perception of affective prosody) depends predominantly on the left inferior frontal gyrus. According to them, comprehension of linguistic prosody requires analysis of the lexical, semantic, and syntactic aspects of pitch modulations. Therefore, the observed pattern of activation seems to indicate that some of these analyses are mediated by the anterior perisylvian language areas. Two fMRI studies [Gandour et al., 2007; Tong et al., 2005], further examined cross-linguistic neural processes involved in the perception of sentence focus (contrastive stress on initial vs. final word) and sentence type (statement vs. question) using discrimination tasks on pairs of stimuli. Compared with passive listening of the same pairs of stimuli, both studies put forward LH lateralized involvement of the supramarginal gyrus and the posterior middle temporal gyrus and RH lateralized involvement of the mid-portion of the middle frontal gyrus extending to the inferior frontal gyrus. For the authors, these results suggest both a leftward asymmetry in temporal and parietal areas in relation to auditory-motor integration as well as working memory demands and higher-level linguistic processing (such as computing the meaning of spoken sentences) and a rightward asymmetry in frontal areas linked to more general auditory attention and working memory processes associated with lower-level pitch processing [Gandour et al., 2007]. Even if they provide important data, these studies do not clearly provide details about cerebral regions involved in the circumscribed processing of prosodic focus (i.e., focus detection), however. Furthermore, both the pairwise comparison tasks used in these studies (which involve strong working memory load and complex comparisons of F_0 patterns) and the fact that the focus conditions were contrasted to rest or to passive listening preclude the interpretation of the results as strictly focus specific.

Finally, a number of Event-Related Potential studies also explored the perception of contrastive focus using congruent vs. mismatching focus patterns [Bögels et al., 2011; Hruska et al., 2001; Johnson et al., 2003; Magne et al., 2005; Toepel et al., 2007]. They put forward processing difficulties (N400 effect) when the linguistic focus structure was incoherent with the actual prosodic realization. In addition, Bögels et al. [2011] suggested that the processing of such a mismatch seems more left lateralized.

This literature review therefore shows that the cerebral network involved in the processing of prosodic focus remains unclear. The goal of this fMRI study is to further clarify the neural processes specifically involved in the auditory perception of prosodic focus using a natural focus detection task (no pairwise comparisons that may involve task-specific demands not limited to the perception of prosody). Two experimental conditions are contrasted, which involve the processing of two types of utterances:

either with narrow prosodic contrastive focus, referred to as "focused" stimuli, or with broad focus (no contrastive focus) referred to as "neutral stimuli." The task is designed so that the participants indirectly pay attention to prosody. The focused and neutral conditions compared with one another are similar in terms of attentional processes. The utterances in both conditions have the same segmental, lexical and syntactic contents and they only differ in their prosodic realization. This paradigm makes it possible to examine processes strictly associated with prosody processing. Note that these processes are of course not independent of semantic and syntactic processes for example.

The comparison between the focused and neutral conditions allows us to test predictions made by the various neural models of prosody presented earlier. If the functional lateralization hypothesis is correct, then, because both conditions involve a linguistic task, there should be no difference in lateralization between the two conditions, which should be both left lateralized. If the acoustic cue lateralization hypothesis is correct, a difference should be observed between the two conditions, as the focused items are associated with increased values of some acoustic cues (such as F_0 and duration). As explained earlier, it has been claimed that the processing of pitch information (and more specifically pitch contour, that spread over long temporal windows) is associated with RH activation. This view is debated; one argument being that the lateralization also depends on the task. By conducting a parametric analysis using a parameter reflecting pitch (F_0) variations specifically related to focus, we can further test the hypothesis that lateralization may reflect hemispheric specialization in the processing of specific acoustic cues, and typically of pitch. This analysis aims at examining the processing of focus-related variations of F_0 among all prosodic parameters (F_0, intensity, duration, phrasing, etc.). If the acoustic cue lateralization hypothesis is correct, then the parametric analysis should yield a RH asymmetry. On the contrary, if the view that cue lateralization depends on the task, and because linguistic tasks are considered as left-lateralized, then the parametric analysis on this focus-related parameter should not necessarily yield a rightward asymmetry and could even reveal a LH bias. Finally, if the AST hypothesis is correct, then both conditions should yield activity in the RH, as the search for a focused item requires temporal integration over a large window.

MATERIAL AND METHODS

Participants

Twenty four healthy adults participated in the experiment (12 females/males, mean age ± SD: 27 years ± 4, age range: 19-34 years). All participants were right-handed according to the Edinburgh Handedness Inventory [Oldfield, 1971], were native speakers of French and had no history of language, neurological and/or psychiatric disor-

ders. They gave their informed written consent to participate in the experiment and the study was approved by the local ethics committee (CPP no. 09-CHUG-14, April 6, 2009).

Stimuli

A corpus of 24 French sentences (Appendix) was recorded by a French female speaker in a soundproof room. All sentences had the same syntactic and syllabic structure: Subject (S: bisyllabic first name) – Verb (V: bisyllabic past tense verb) – Object (O: monosyllabic determiner + bisyllabic noun), as in the following example: "Marie serrait la poupée" ("Marie was holding tight onto the doll").

To avoid the predictability of focus location (discussed later), all 24 sentences were recorded in two ways: with narrow contrastive focus on the subject (e.g.,"MARIE serrait la poupée") and with narrow contrastive focus on the object (e.g., "Marie serrait la POUPEE"). For control, the same 24 sentences were recorded with a neutral statement intonation, i.e., with broad focus (e.g., "Marie serrait la poupée").

From these sentences, we contrasted two experimental conditions: narrow contrastive Focus (F) and Neutral (N, control). The direct comparison F vs. N allowed identifying cerebral regions specifically involved in the detection of narrow contrastive focus. The F condition consisted of the 2 × 24 sentences with narrow contrastive focus, each presented once (48 sentences total). The N condition consisted of the 24 neutral (broad-focused) sentences, each presented twice (48 sentences total).

All the recorded utterances were digitized at a sampling rate of 44.1 kHz with a 16-bit quantization. An acoustic and perceptual validation was conducted to check that prosodic contrastive focus had indeed been produced for the narrow focused utterances and not for the broad focused (neutral) utterances. All narrow focused stimuli displayed the typical acoustic correlates described in Dohen and Lœvenbruck [2004]. Figure 1 provides examples of recorded stimuli in a neutral and a narrow focused (subject focus) case for the same sentence ("Marie serrait la poupée"). In this example, we can see that pitch is typically higher on the subject when it is focused (neutral: max: ∼ 200 Hz; focus: max: ∼ 220 Hz). It also shows that after focus (on the verb and object of the utterance), pitch is typically lower and flatter compared with the neutral utterance.

F_0-Related Acoustic Parameter

As discussed previously, hemispheric lateralization may reflect hemispheric specialization in the processing of specific acoustic cues, such as duration or fundamental frequency (pitch, F_0). As pitch variations have been shown to be an important acoustic correlate of contrastive prosodic
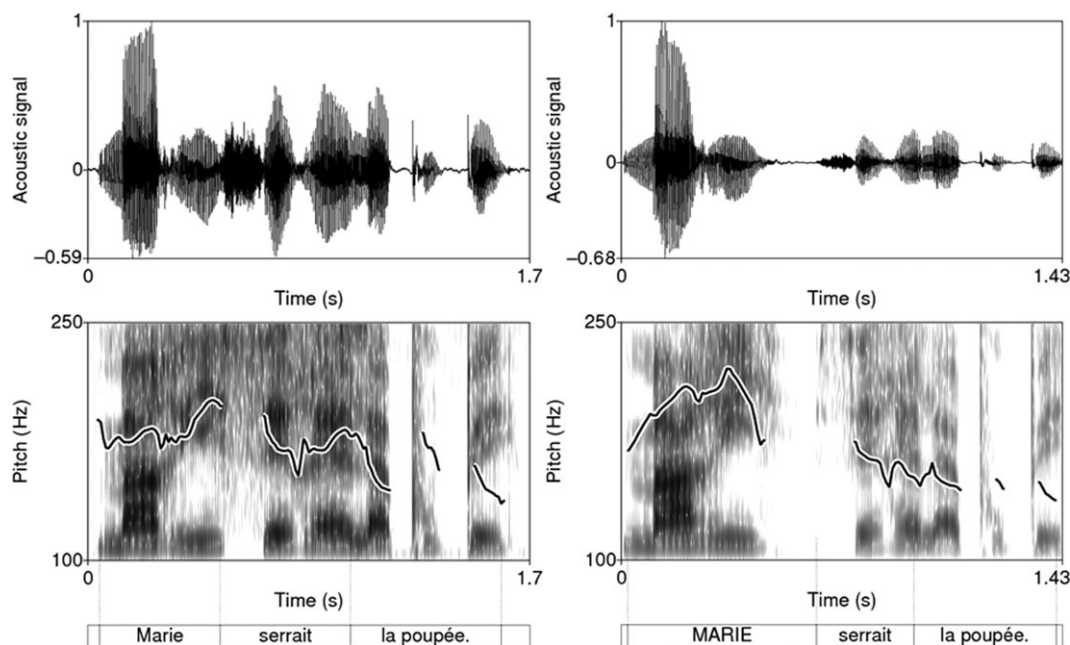
Figure 1.

Examples of recorded stimuli in a neutral and a focused (subject focus) case for the same sentence ("Marie serrait la poupée", "Marie was holding tight onto the doll"). As shown in the spectrogram (lower figure), pitch is higher on the subject when it is focused (neutral: max: $\sim$ 200 Hz; focus: max: $\sim$ 220 Hz) and, after focus (on the verb and object of the utterance), pitch is lower and flatter compared with the neutral utterance.

focus [Dohen and Lœvenbruck, 2004], it seemed relevant to carry out a parametric analysis using an acoustic cue reflecting F_0 variations specifically related to focus.

In French, the fundamental frequency peak on the focused constituent is the focal accent and is a strong cue to the listener as to which constituent was focused (Fig. 1, right panel: F_0 peak on the focused item "MARIE"). The maximum of F_0 over the utterance could therefore seem an appropriate acoustic parameter reflecting F_0 variations associated with focus. The perception of focus is relative to the utterance as a whole, however, the level of the F_0 peak corresponding to the focal accent is processed relative to the F_0 contour of the utterance, taking declination into account. Declination is a general lowering (decline) of pitch throughout declarative sentences (e.g., in French: Delgutte, 1978]. As Ladd [1996] puts is, it means that "a pitch movement at the beginning of a phrase will be higher than the same pitch movement later in the phrase" (as an illustration, Fig. 1, left panel: the initial F_0 value in this neutral declarative sentence is higher than the final F_0 value). It is however well known that listeners compensate for this declination [Liberman and Pierrehumbert, 1984] and have no difficulties in identifying a focal accent at the end of an utterance even if the corresponding F_0 peak is actually lower than those at the utterance beginning. To take declination effects into account, the acoustic parameter was computed relative to the declining F_0 backdrop over the utterance. Our corpus consisted of subject (S) focused utterances and object (O) focused utterances. Declination results in the focal accent on the O being sometimes smaller than the F_0 peak observed on the unfocused S of the same utterance. The declination was compensated for by computing a declination factor for each Neutral rendition of the stimuli [αdec = max F_0(S)/max F_0(O)] and choosing the minimal declination factor as the speaker's declination factor. All O F_0 peaks were then multiplied by this factor. Because the speaker can have varying F_0 ranges from one stimulus to another and because listeners also compensate for this by detecting focal accents relative to the F_0 range of a given utterance, we compensated for interutterance variations by normalizing all the F_0 values relative to the minimum F_0 (0 after normalization) and maximum F_0 (1 after normalization) produced by the speaker. After these two manipulations, the declination-corrected F_0 range was computed as the difference between the normalized F_0 maximum and minimum for each stimulus. We refer to this parameter as the DCF_0R. We checked that this DCF_0R parameter correctly reflected focus and used it in the parametric analysis. Table I shows the mean values of the DCF_0R parameter for Neutral and Focused stimuli, as well as for Subject-Focused and Object-Focused stimuli. DCF_0R values are significantly lower for Neutral items than for Focused items [$F(1,94) = 161.78$, $P < 0.001$]. Furthermore, the values are significantly different for the

TABLE I. Mean values and standard deviation of the declination-corrected range of F_0 (DCF_0R) parameter in the different conditions

	Mean	SD
Neutral	0.69	0.02
Focused	0.82	0.08
Subject-focus	0.78	0.04
Object-focus	0.87	0.08

three different prosodic conditions [$F(2,93) = 132.50$, $P < 0.001$]. DCF_0R is lower for Neutral items than for Subject-Focused items, and lower for Subject- than Object-Focused items.

Tasks

During the experiment, participants were indirectly asked to judge whether each utterance they heard contained contrastive focus (either on the subject or the object) or not. In order for the task to be natural, the participants were not directly asked to detect prosodic focus. They were told that they would hear utterances extracted from a dialogue in which a first speaker (S1) uttered a sentence, then a second speaker (S2), believing he had misunderstood part of the sentence, questioned S1 by repeating the sentence he had understood in a question mode, and finally S1 repeated the first sentence correcting what S2 had misunderstood. For example: S1: "Marie serrait la poupée.", S2: "Sarah serrait la poupée", S1: "MARIE serrait la poupée".

The participants were told that they would only hear the second utterance produced by S1 and that there were two possible dialogue situations:

Situation A: S2 misunderstood part of the sentence and S1 corrected what had been misunderstood (narrow contrastive focus case);

Situation B: S2 understood well and S1 repeated the sentence in a neutral mode (broad focus case).

The task was to identify the dialogue situation (A or B) from hearing the second utterance produced by S1. To perform the task, participants were instructed to judge whether the utterance contained or not correction.

The "Yes" (there is a correction) and "No" (there is no correction) responses were provided with the index and the middle fingers of the right hand, by means of two response keys. They were recorded and the performance of task execution was evaluated. Participants were trained to the task outside the scanner before the actual experiment. The training sentences were different from those presented during the fMRI experiment. Each response was followed by a feedback in order to inform the participants about their performance. The fMRI session began only when the experimenter was sure that the participants had clearly understood (did not make any errors) the task and were able to perform it.

Functional MRI Paradigm

The stimuli were presented via E-Prime (E-prime Psychology Software Tools Inc., Pittsburgh, USA) running on a PC computer. They were delivered by means of MRI-compatible electrostatic headphones to the rate of one every 2 seconds. Between stimuli, the subjects were asked to maintain their gaze on a white fixation cross displayed at the center of a black screen. A total of 96 stimuli were presented: 48 utterances with contrastive focus, (24 subject-focused and 24 object-focused utterances) and 48 broad-focused (Neutral) utterances. Each sentence from the corpus was thus presented twice in each condition: once with focus on the subject and once with focus on the object in the Focus condition and twice with broad focus on the Neutral condition. A pseudo-randomized event-related fMRI paradigm, including one functional run, was designed based on the optimization of the onset for each type of stimuli for each condition [Friston et al., 1999]. The functional run included 48 events per prosodic condition (Focus, Neutral) and 30 null-events to provide an appropriate baseline measure [Friston et al., 1999]. The null-events consisted of a fixation cross presented at the centre of the screen. The average inter-stimulus interval was 4s and the functional run lasted approximately 9 min. To stabilize the magnetic field in the MRI scanner, five "dummy" scans were added at the beginning of the run and were removed from the analyses. Overall, 168 functional volumes were acquired.

MR Acquisition

The fMRI data were acquired using a whole-body 3T scanner (Bruker MedSpec S300). Functional images were obtained using a T2*-weighted, gradient-echo, echoplanar imaging (EPI) sequence with whole-brain coverage (TR = 3 s, spin echo time = 40 ms, flip angle = 77°). Each functional scan comprised 39 axial slices parallel to the antero-posterior commissural plane acquired in interleaved order (3×3 mm^2 in plane resolution with a slice thickness of 3.5 mm). Images were corrected for geometric distortions using a B0 fieldmap. In addition, a high-resolution T1-weighted whole-brain structural image was acquired for each participant (MP-RAGE, volume of $256 \times 224 \times 176$ mm^3 with a resolution of $1.33 \times 1.75 \times 1.37$ mm^3).

fMRI Data Processing

Data were analyzed using the SPM5 software package (Wellcome Department of Imaging Neuroscience, Institute of Neurology, London, UK) running on Matlab 7.1 (Mathworks, Natick, MA, USA). The ROI analysis was performed using the MarsBar software (http://marsbar.sourceforge.net/). The brain regions involved in the different contrasts, we examined were labeled using a macroscopic parcellation of the MNI single subject reference brain [Tzourio-Mazoyer et al., 2002]. For

visualization, the statistical maps were superimposed on a standard brain template using the MRICRON software (http://www.sph.sc.edu/comd/rorden/mricron/).

Spatial pre-processing

For each participant, the functional images were time-corrected (slice timing). All volumes were then realigned to correct for head motion. Unwrapping was performed using the individually acquired fieldmaps to correct for interaction between head movements and EPI distortions [Andersson et al., 2001]. The T1-weighted anatomical volume was coregistered to mean images created by the realignment procedure and was normalized to the MNI space using a trilinear interpolation. The anatomical normalization parameters were then used for the normalization of functional volumes. All functional images were then smoothed using a 6-mm full-width at half maximum Gaussian kernel to improve the signal-to-noise ratio and to compensate for the anatomical variability among individual brains.

Statistical analyses

For each participant, the two conditions of interest (F: Focus, N: Neutral) were modeled using the General Linear Model [Friston et al., 1995]. The six realignment parameters were also included as covariates of no interest. The blood-oxygen-level dependence response for each event was modeled using a canonical hemodynamic response function (HRF). Before estimation, a high-pass filtering with a cutoff period of 128 s was applied. Beta weights associated with the modeled HRF responses were then computed to fit the observed blood-oxygen-level dependence signal time course in each voxel for each condition (F and N). Individual statistical maps were calculated for each condition contrasted with the related baseline and subsequently used for group statistics. To draw population-based inferences [Friston et al., 1998], a second-level random effect group analysis was carried-out. One-sample t-tests were calculated to determine brain regions specifically involved in F vs. N contrasts. The results are reported at a False-Discovery Rate corrected level (FDR) of $P < 0.05$ (cluster extent of 25 voxels).

To investigate the degree of hemisphere predominance, we defined several Regions of Interest (ROI) based on whole-brain activation and obtained by contrasting F vs. N. Specifically, we retained all the activated voxels included within 5 mm radius around each peak of activation (Table I), in left and RHs. The MarsBar software (http://marsbar.sourceforge.net/) was used to build ROIs. For each pair (left-right) of ROI and each participant, the % of MR signal change was extracted and the values were entered in an analysis of variance (ANOVA) with "Hemisphere" (Left, Right) and "Prosodic Condition" (F, N) as within-subject factors. We predicted that differential involvement of hemispheres may be reflected by a signifi-

cant main effect of Hemisphere according to this analysis. Moreover, a significant interaction Hemisphere x Prosodic Condition may reflect the modulation of hemisphere activity by the prosodic condition.

Finally, an item-wise parametric analysis [Buchel et al., 1996, 1998] was conducted to identify brain regions specifically modulated by the DCF_0R acoustic parameter. For each participant, the Focus and Neutral conditions were collapsed and modeled as a unique regressor in a second General Linear Model. As previously, this regressor was convolved with the canonical HRF and the six realignment parameters were inserted as covariates of no interest. To examine parametric activity modulation on DCF_0R values, a parametric regressor was added for each stimulus. A contrast image representing the parametric modulation on DCF_0R was then computed at the first level for each participant. Based on the contrast images, a random-effect group analysis was then performed using a one-sample t-test. The results are reported at a FDR corrected level of $P < 0.05$ (cluster extent of 25 voxels).

RESULTS

Behavioral Results

Behavioral responses recorded during the fMRI experiment showed that participants performed the task correctly for both prosodic conditions: Focus ($M = 93.92\%$, SD = 5.63%) and Neutral ($M = 98.35\%$, SD = 3.65%). The correct response level was significantly above chance (50%) for both Focus (t (23) = 37.35, $P < 0.001$) and Neutral (t (23) = 63.40, $P < 0.001$) conditions. Moreover, the participants were more accurate in the Neutral than in the Focus condition (t (23) = 4.27, $P < 0.001$).

fMRI Results

Cerebral network specifically involved in F vs. N

The contrast F vs. N (Table II, and Fig. 2) revealed the network of regions which were significantly more involved in Focus than in Neutral condition: left inferior orbital part (F3O) of inferior frontal gyrus, right inferior triangular part (F3T) of inferior frontal gyrus, bilateral inferior parietal regions (P2), left superior parietal region (P1), right T1, and bilateral T2.

#### ROI analysis on hemispheric differences

The ROI analysis (Table III, Fig. 3) showed a significant main effect of the hemisphere for the superior frontal [SMA; F (1,22) = 6.14, $P < 0.05$], superior parietal [P1; F (1,22) = 5.27, $P < 0.05$] and superior temporal [T1; F (1,22) = 19.5; $P < 0.05$] gyri. The SMA and P1 were significantly lateralized to the right for both prosodic conditions whereas T1 was lateralized to the left. More interestingly, we found a significant interaction between Hemisphere

TABLE II. Activation peaks provided by the random-effect group analysis for Focus vs. Neutral contrast (P < 0.05, FDR corrected)

Condition	Lobe	Anatomical description	aal-Label	H	x	y	z	T
F > N	Frontal	Inferior frontal gyrus, pars orbitalis	F3O	R	54	30	−4	5.63
		Inferior frontal gyrus, pars triangularis	F3T	L	−51	18	7	5.41
		Supplementary motor area	SMA	L/R	±3	33	42	4.93
	Parietal	Inferior parietal gyrus	P2	L	−51	−51	56	5.42
			P2	R	51	51	53	4.65
		Superior parietal gyrus	P1	L	−42	−72	53	5.34
	Temporal	Superior temporal gyrus	T1	R	63	−42	14	4.83
Middle temporal gyrus			T2	L	−51	−33	−7	4.61
			T2	R	51	−36	4	4.78

Cluster extend threshold of 25 voxels, MNI coordinates. L/RH, left/right hemisphere, labeling according to Tzourio-Mazoyer et al., 2002.

and Prosodic Condition for the superior temporal gyrus [T1; *F* (1,22) = 12.81; *P* < 0.05] and a marginally significant interaction for the triangular part of the IFG [F3T; *F* (1,22) = 3.60; *P* = 0.07]. This suggests that the modulation of the activity by the prosodic condition is different across conditions and hemisphere (Fig. 4). Right T1 appears to be more
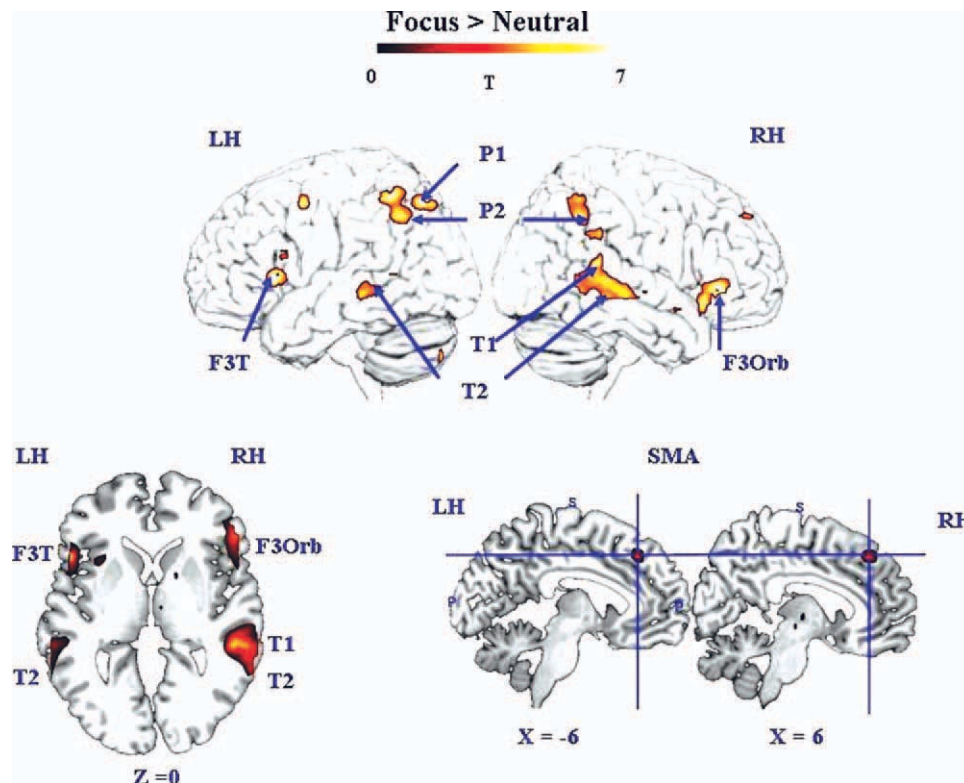
Figure 2.

Cerebral activity involved in the F vs. N contrast (random-effect group analyses, *P* < 0.05, FDR corrected, cluster extend threshold of 25 voxels, LH/RH: left/right hemisphere). The activation is projected onto surface rendering (upper row) and onto 2D anatomical slices in axial and sagittal orientation (lower row) represented in neurological convention (left is left hemisphere). The activated regions are indicated by using blue arrows. The MNI coordinates of the activation are also mentioned. LH, left hemisphere, RH right hemispheres; F3Orb, inferior frontal gyrus, pars orbitalis; F3T, inferior frontal gyrus, pars triangularis; SMA, supplementary motor area; P2, inferior parietal gyrus; P1, superior parietal gyrus; T1, superior temporal gyrus; T2, middle temporal gyrus. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

TABLE III. Anatomical Description and Labels of ROI

Anatomical description	Label	Coordinates			Main H effect		Interaction effect	
		x	y	z	F	P	F	P
Supplementary motor area	SMA	±3	33	42	6.14	0.02	0.21	0.65
Superior parietal gyrus	P1	±42	−72	53	5.27	0.03	0.56	0.46
Middle temporal gryus	T2	±51	−33	−7	0.32	0.57	1.25	0.28
Superior temporal gyrus	T1	±63	−42	14	19.50	0.0001	12.81	0.001
Inferior parietal gyrus	P2	±51	−51	56	0.27	0.60	1.87	0.18
Inferior frontal gyrus (pars triangularis)	F3T	±51	18	7	0.07	0.78	3.60	0.07
Inferior frontal gyrus (pars orbitalis)	F3O	±53	30	−4	0.84	0.36	0.77	0.39

The MNI coordinates indicate the activation peak of ROI. P and F values show the significance of the main effect for the hemisphere and for the interaction hemisphere (left, right) X Prosodic condition (focus, neutral).

involved in the F than in the N condition while its intensity of signal change in the LH is the same for both prosodic conditions. F3T is always more involved for F than for N but this effect is suggested to be stronger in the LH.

SPM parametric analysis with DCF₀R

The item-wise parametric analysis on DCF_0R values revealed the network of cerebral regions positively correlated with the DCF_0R parameter (Table IV and Fig. 5). These regions included the bilateral IFG (left triangular and right orbital), superior medial frontal gyrus (SMA),
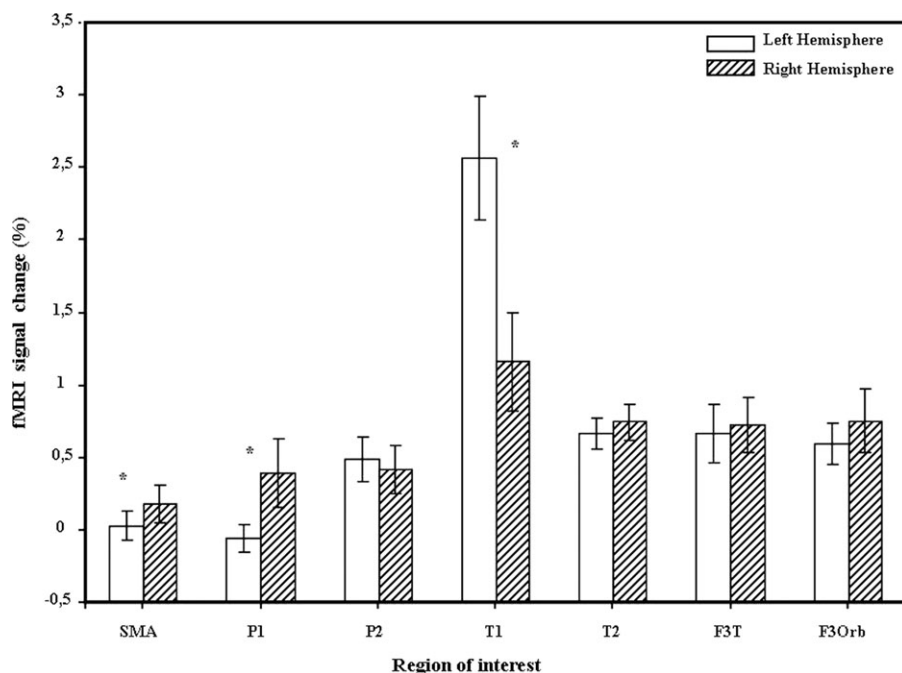
posterior part of the superior temporal gyrus as well as the left precentral and inferior parietal gyri and the bilateral anterior insula, left cerebellum and right thalamus.

DISCUSSION

This study examined the neural processes involved in the auditory perception of prosodic focus using a natural focus detection task. Participants were placed in a natural dialogue situation, in which they simply had to judge whether the speaker had performed a correction in her utterance or not. The task was thus to indirectly decide

Figure 3.
Brain regions showing inter-hemispheric difference with ROI analysis (for each region, the % of MR signal is indicated). SMA, supplementary motor area; P1, superior parietal gyrus; P2, inferior parietal gyrus; T1, superior temporal gyrus; T2, middle temporal gyrus; F3T, inferior frontal gyrus, pars triangularis; F3Orb, inferior frontal gyrus, pars orbitalis.
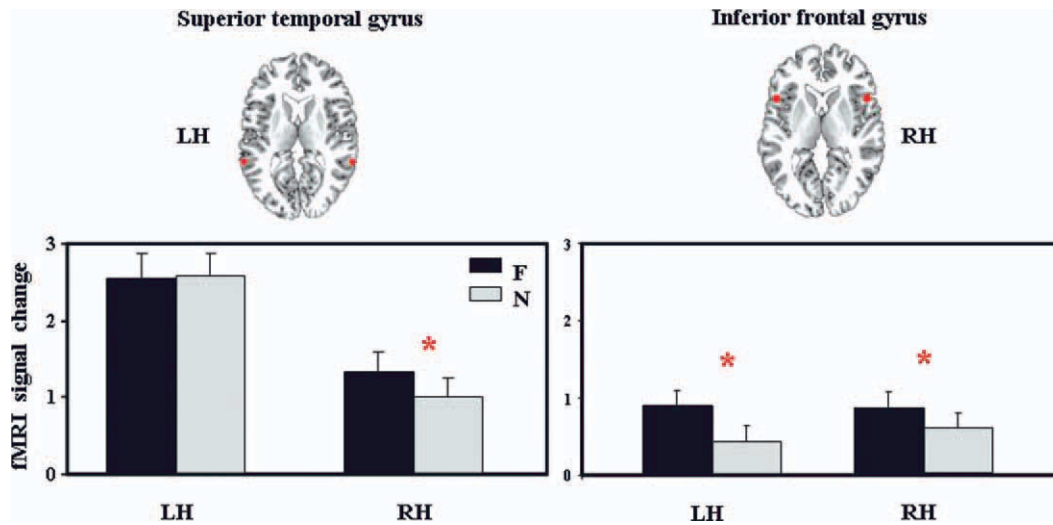
Figure 4.

Brain regions showing interaction between hemisphere and prosodic condition (focus and neutral) in the ROI analysis (for each region, the percentage of signal change is indicated, LH/RH, left/right hemisphere; F, focus condition; N, neutral condition). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

whether an utterance contained contrastive prosodic focus. As the focused and neutral stimuli only differed in their prosodic realization and contained exactly the same segmental and lexical content, the differences between conditions cannot be due to lexical or syntactic changes.

The direct comparison F vs. N sheds light on the cerebral network specifically recruited for the processing of prosodic focus. The parametric analysis performed with the DCF_0R as parameter provides information for discussing the effect of this specific prosodic cue on the activity of bilateral brain regions.

Perceiving Prosodic Focus: F vs. N Contrast

We discuss the brain involvement data in the F vs. N contrast for temporal, parietal, and frontal cortices. This contrast captures cerebral regions specifically involved in prosodic

focus detection relative to the processing of neutral prosodic cues. We discuss whether the pattern of cerebral regions associated with the detection of prosodic focus is compatible with a dual-stream model. Dual-stream models of speech perception, inspired by neural models of vision, posit the existence of two pathways, a ventral ("what") pathway including temporal regions, and a dorsal ("how") pathway including parietal and frontal regions [e.g. Hickok and Poeppel, 2000, 2004, 2007; Rauschecker, 2011; Rauschecker and Scott, 2009; Rauschecker and Tian, 2000; Scott and Johnsrude, 2003]. In addition, to test the predictions of several neural model of prosody, we examine lateralization effects.

Temporal cortex

The superior temporal gyrus was activated bilaterally in both prosodic conditions. As we explained in the results

TABLE IV. Activation Provided by the Random-Effect Group Parametric Analysis on F_0 ($P < 0.05$, FDR Corrected, Cluster Extend Threshold of 25 Voxels, MNI Coordinates, L/RH: Left/Right Hemisphere)

Lobe	Anatomical description	aal-Label	H	x	y	z	T
Frontal	Inferior frontal gyrus, pars triangularis	F3T	L	−51	18	7	7.07
	Inferior frontal gyrus, pars orbitalis	F3O	R	51	33	−4	7.86
	Supplementary motor area	SMA	L/R	±0	33	46	6.60
	Precentral gyrus	PRE	L	−45	3	56	5.64
	Insula	IN	L	−39	21	3	3.63
		IN	R	39	24	−7	4.60
Temporal	Superior temporal gyrus	T1	L	−54	−21	0	5.24
		T1	R	60	−33	4	5.90
Parietal	Inferior parietal gyrus	P2	L	−48	−57	42	6.56
Cerebellum			L	−18	−81	−39	5.06
Subcortical	Thalamus	THA	R	6	−12	−4	4.31
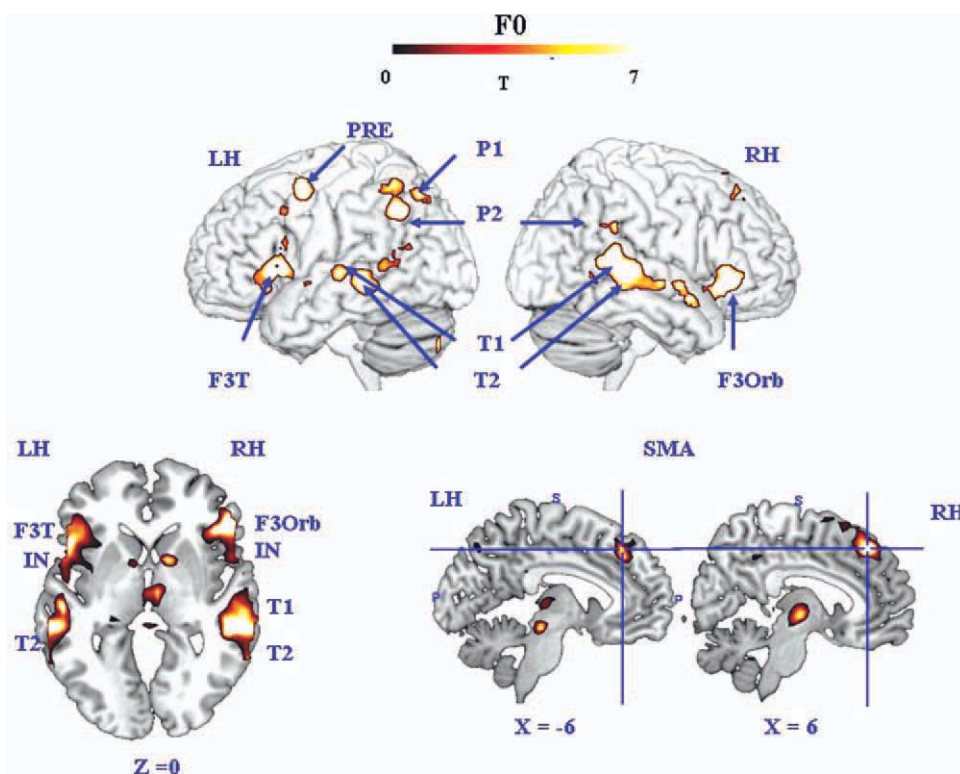
Figure 5.

Surface rendering of brain regions involved in the F_0 parametric analysis (random-effect group analyses, $P < 0.05$, FDR corrected, cluster extend threshold of 25 voxels, LH/RH: left/right hemisphere). Activated regions are indicated by blue arrows. LH, left hemisphere; RH, right hemispheres; F3Orb, inferior frontal gyrus, pars orbitalis; F3T, inferior frontal gyrus, pars triangularis; SMA, supplementary motor area; P2, inferior parietal gyrus; P1, superior parietal gyrus; T1, superior temporal gyrus; T2, middle temporal gyrus. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

section, the ROI analysis on the posterior part of the superior temporal gyrus showed that LH activation is predominant for both conditions. This result is in agreement with the notion of LH predominance for linguistic components (i.e., phonological, lexical and syntactic and structures, shared by both prosodic and neutral conditions), as suggested by Vigneau et al. [2011, 2006]. Furthermore, the right superior temporal gyrus activity was modulated by the prosodic condition. A right dominant activity during prosody perception has been associated with prosodic cues that need to be integrated over large temporal windows. Studies on global intonation contours confirm the role of the right temporal lobe in the processing of slowly changing prosodic cues [Meyer et al., 2004; Zatorre and Gandour, 2008]. However, in this study, both conditions involved the processing of slowly changing acoustic cues (both utterances display F0 variations which are known to be perceived and processed by listeners) and the rightward asymmetry is observed for the contrast between these conditions. The difference lies in the fact that in the Focus condition the processing of the acoustic cues will lead to the detection of prosodic focus, which is not the case in the Neutral condition. The participants were trying to detect specific prosodic cues in the stimuli (focal accent, lengthening, post-focal deaccentuation, etc., c.f. Dohen and Lœvenbruck, 2004]. This expectation was met in the Focus condition only. The rightward modulation of the activity in the STG would then be linked to the actual detection of expected slowly varying prosodic acoustic cues, but not just to the processing of slowly varying cues because this was involved in both conditions. Added to the fact that the LH is dominant for both conditions, this implies that the Asymmetric Sampling in Time hypothesis needs to be refined to take our results into account. The theoretical account provided by von Kriegstein et al. [2010] can be extended to our results. This accounts aims at explaining the bilateral involvement of regions in superior temporal gyrus/superior temporal sulcus (STG/STS) for changes in speaker-related vocal tract parameters as well as for speech recognition. These authors speculate that their results can be interpreted in terms of distinct but coupled mechanisms in the left and RH. RH computation of information changing at slower time scales (such as vocal tract length of the speaker) can help predict information at

faster time scales (such as formant positions determining phonemes of that speaker), processed in the LH. According to the authors, the right posterior STG/STS activation reflects the extraction of speaker parameters, which is used by an internal model to recognize the message. This view is in agreement with the view that speech perception proceeds on the basis of internal forward models [Poeppel et al., 2008, see also the discussion on perception-action interaction below). The internal model is hypothesized to be updated regularly with the results of the computations at the segmental (short resolution) and syllabic (longer resolution) levels. In our study, the increase of right STG activity in the Focused condition compared with the Neutral condition can be interpreted in terms of increased connectivity with the LH to check that the prosodic variations can be linguistically recognized as corrective focus. We speculate that the results of the computations in the right STG are fed to an internal model used by the left STG to better predict the linguistic phenomenon of focus.

The middle temporal gyrus (T2) was bilaterally involved in the Focus condition (compared with the Neutral condition). This region, typically in the LH, is associated with lexical [Lau et al. 2008] and high-level semantic [Démonet et al., 2005] processing. According to the dual-stream model of functional anatomy of language perception [Hickok and Poeppel, 2007], the ventral auditory stream (including middle temporal gyrus) is involved in processing speech signals for comprehension. The ventral stream maps sensory and phonological representations onto lexical conceptual representations (i.e., sound-to-meaning). The activation of the middle temporal gyrus observed in the detection of prosodic focus is in agreement with this model, as the presence of prosodic focus on a word highlights it and probably increases its lexical processing. It can be hypothesized that the detection of prosodic focus activates semantic processing, as prosodic structure interpretation, and typically focus projection, is bound with semantic interpretation [e.g., Welby, 2003]. The fact that we found a bilateral involvement of the middle temporal gyrus whereas a RH bias was found for the STG can be interpreted by processing level. At a more lexical level, the right dominance would not be as strong, longer timescale computations being less crucial. This bilaterality can still be interpreted as the manifestation of an interplay between the two hemispheres with RH computations of slowly changing prosodic information serving to predict the lexical-level computations carried out in the LH, but being less critical.

Parietal cortex

Prosodic focus detection, relative to neutral sentence processing, involved the bilateral inferior parietal lobule. The involvement of the left side of the inferior parietal lobule is consistent with previous observations during speech processing tasks. It has been suggested that this associative area is an interface between sound-based repre-

sentations of speech in auditory cortex with articulatory-based representations of speech in frontal cortex [Hickok and Poeppel, 2000, 2007; Rauschecker and Scott, 200]. In analogy with the dorsal pathway hypothesized in vision, the left inferior parietal cortex would play a role within a temporo-parieto-frontal network functioning as an interface system between auditory and motor processes. It seems reasonable to hypothesize that an analogous interface region could be involved in the RH to process the prosodic cues to focus. Although the involvement of the RH in prosodic processing is suggested by Hickok and Poeppel [2007] on the basis of several results in the literature, some of which are cited earlier (c.f., temporal cortex section), it is not clear whether the RH involvement also applies to the inferior parietal lobule. Hesling et al. [2005] also found right and left supramarginal gyrus association in the processing of expressive speech, involving a high degree of prosodic expression. They relate bilateral involvement to specific processing differing in window length and in acoustic cues decoding in accordance with the AST hypothesis. Other studies of prosodic processing have found varying results. Tong et al. [2005] found stronger hemodynamic response of left than right supramarginal gyrus, but their control task was a passive listening task with the same stimuli as those used in the prosodic task. Therefore, it can be argued that the control task partly involved the same pitch information decoding as the prosodic task, which may have hidden the RH involvement.

In addition, we found a LH response of the superior parietal lobule. Although the involvement of this region is rarely mentioned in studies of prosodic processing, Lœvenbruck et al. [2009] also reported a recruitment of the left superior parietal lobule in a prosodic focus production task. In that study, the left superior parietal lobule was activated in three pointing tasks (pointing with the finger, pointing with the eye, or gazing, and pointing with the voice, or producing prosodic focus). These results together with the present result on prosodic focus detection seem to suggest that the left superior parietal lobule may be recruited in the processing of pointing generally and of prosodic pointing or focus more specifically.

Frontal cortex

The F vs. N contrast increases the hemodynamic response bilaterally in the inferior frontal region, specifically in the right orbital and left triangular part of this region. As concerns the orbital part, Meyer et al. [2002] found that prosodic sentence compared with syntactic sentence processing, induces a stronger hemodynamic response of the right inferior frontal region and specifically of its opercular part. They speculated that right fronto-opercular cortex would be involved in the extraction of slow pitch information. Although the peak of the activation measured in their study ($x = 41$, $y = 10$, $z = 12$) was not exactly the same as in ours ($x = 54$, $y = 30$, $z = -4$),

we could extend their speculation to the orbital part of the right inferior frontal region and consider that this region may reflect the processing of slow speech modulation in our results. However, as explained earlier, both conditions involved this type of processing. An alternative hypothesis could be that this RH frontal activation is part of a RH dominant network involved in the processing and detection of some slow-varying acoustic cues, based on an internal forward model making predictions. Based on suggestions by Rauschecker and Scott [2009], we speculate that the RH frontal activation could reflect the involvement of an emulator or internal model, which may provide predictions facilitating the focus detection process. We will come back to this point in the discussion on the perception-action interaction.

In addition to the right pars orbitalis, we found activation of Broca's area, or, more precisely, of the left pars triangularis. The ROI analysis on this region suggested that the activity is modulated by the prosodic condition in both hemispheres. But the marginal significant interaction observed in this region, suggests that this modulation is more important in the left than in the RH. Broca's area has been considered to be involved in complex syntactic processing when thematic role monitoring is required, i.e., the processing of "who-does-what-to-whom" [Caplan et al., 2000; Just et al., 1996]. These studies have shown the involvement of Broca's area in plausibility judgments about syntactically complex constructions (with cleft-object sentences, or sentences with center-embedded clauses), the latter requiring intricate tracking of thematic roles. Studies of the comprehension deficits in aphasic patients also suggest that Broca's area is involved in thematic role processing [e.g., Caplan et al., 1985; Friederici and Gorrell, 1998; Rigalleau et al., 2004; Schwartz et al., 1987]. Broca's area has also been shown to be involved in the production of sentences requiring thematic role processing [Caplan and Hanna, 1998; Collina et al. 2001, Webster et al. 2001; Whitworth, 1995]. In addition, the left inferior frontal gyrus was shown to be involved in a study of the production of prosodic focus and syntactic extraction in French [Lœvenbruck et al., 2005]. This result was interpreted by the authors as a strong involvement of Broca's region in the monitoring of the agent of an action, through either prosody or syntax.

In this study, the utterances to be processed involved prosodic focus on either the agent (the subject) or the patient (the object) of the action. The involvement of the left inferior frontal gyrus during the processing of these utterances together with the earlier-mentioned literature review therefore suggest that the left inferior frontal region is a parser of action structure, necessary in both production and perception. In the framework of a dual-stream model of speech perception, as mentioned earlier about the involvement of the right inferior frontal region, the left inferior frontal activation can be interpreted as evidence for the implication of an internal forward model used to predict sensory outcomes which are to be compared with actual stimuli in the process of focus detection.

The supplementary motor area, SMA, was bilaterally involved when Focus was compared with Neutral condition. Moreover, the ROI analysis suggests stronger hemodynamic response of the right side of the SMA for both prosodic conditions relative to rest but no difference in the modulation of the activity by the prosodic condition between hemispheres. As suggested by Kotz and Schwartze [2010], during speech perception the auditory information is transmitted to the frontal cortex in order to be integrated within the temporal event structure. The auditory information goes through successive structures: cerebellum, thalamus and SMA. This result highlights the interaction between speech perception and motor processes. Furthermore, the involvement of the SMA observed in our results is consistent with the finding of Geiser et al. [2008] who found SMA activation during rhythm processing and suggested that the role of the SMA would be to process "acoustically marked temporal intervals". Prosodic focus indeed has marked durational correlates.

Our study shows the cerebral network involved in the detection of contrastive prosodic focus. Our results are in agreement with a model of speech perception involving ventral as well as dorsal regions (referred to as the dual-stream model). In addition to the left superior temporal gyrus, we found ventral activations including the right superior temporal gyrus and the bilateral middle temporal gyrus and dorsal activations including the bilateral inferior parietal gyrus, the left superior parietal gyrus and the bilateral inferior frontal gyrus. The stronger involvement of the RH could be explained by the processing of slowly changing acoustic information crucial in prosodic focus detection as well as the fact that expected prosodic cues (participants expected to find acoustic cues to prosodic focus or not and specifically paid attention to them) are actually perceived which is the case in the F but not in the N condition.

Interpretation in the Framework of Dual-Stream Models and Perception-Action Interaction

The observation of activations in temporal cortex as well in parietal and inferior frontal regions is consistent with dual-stream models of speech perception and can be interpreted as evidence for a perception-action interaction during prosodic focus detection. Such an interaction is discussed in many recent works on speech perception or comprehension [Pickering and Garrod, 2007; Sato et al., 2009; Schwartz et al., 2010; Skipper et al., 2005].

In dual-stream models, two pathways can be involved in speech perception, a ventral pathway involved in lexical and semantic processing and implying the superior and middle portions of the temporal lobe, and a dorsal stream linking auditory speech representations in the auditory cortex and articulatory representations in the ventral

premotor cortex (vPM) and the posterior part of the inferior frontal gyrus (pIFG), with sensorimotor interaction interfaced in the supramarginal gyrus [SMG; Rauschecker, 2011; Rauschecker and Scott, 2009] or in area SPT [a brain region within the planum temporale near the parieto-temporal junction; Hickock and Poeppel, 2007].

Hickok and Poeppel [2000, 2004, 2007] argue that this auditory-motor interaction is crucial in speech development in children and in learning new verbal forms (which might necessitate more sensory guidance than known forms) in adults. In a more recent proposal, Hickok et al. [2011] have included a state feedback control architecture for both speech production and perception. This involves an internal model that makes forward predictions about both the dynamic state of the vocal tract and about the sensory consequences of those states. In short, the sensory feedback control model of speech production includes pathways both for the activation of motor speech systems from sensory input (the feedback correction pathway) and for the activation of auditory speech systems from motor activation (the forward prediction pathway). The same sensory-to-motor feedback circuit can be excited by other's speech, resulting in the activation of the motor speech system from speech listening. According to this view, just as auditory feedback is necessary to generate corrective signals for motor speech acts, others' speech can be used to tune new motor speech patterns. Motor networks are activated during passive speech listening not because they are critical for analyzing phonemic information for perception (a view defended in the Motor Theory of speech perception and its latest developments associated with the discovery of mirror neurons, Fadiga et al., 2009; Rizzolatti and Arbib, 1998] but rather because auditory speech information later benefits production. According to Hickok et al., motor activations observed are not necessarily causally related to speech perception and could be epiphenomenal consequences of associative processes. The forward sensory prediction affords a natural mechanism for a limited role of the motor system on perception.

A somewhat different view of the dual-stream organization has been described by Rauschecker and Scott [2009] and Rauschecker [2011]. According to them, speech perception involves a forward mapping, in which speech is decoded in the anteroventral stream to the inferior frontal cortex and is transformed into motor-articulatory representations. These frontal activations are transmitted to the inferior parietal lobule and posterior superior temporal cortex where they are compared with auditory and other sensory information. In motor control theory terms, this modelling includes forward models that predict the consequences of actions and inverse models that determine the motor commands required to produce a desired outcome. The inferior parietal cortex is thus suggested as an interface for feed-forward information from motor preparatory networks in the prefrontal and premotor cortices to be matched with feedback signals from sensory areas. In this view, predictive motor signals do modify activity in sen-

sory structures and have a more crucial role than the one advocated by Hickock and Poeppel [2000, 2004, 2007].

The observation of stronger activity of a network including temporal, parietal and frontal regions in the Focus condition compared with the Neutral condition, suggests that the detection of prosodic focus is a task for which a dual pathway is involved.

First, our findings support the activation of a bilateral ventral stream. The fact that right-dominant STG activities are observed in the contrast between Focus and Neutral conditions, added to the fact that both conditions involve a left STG dominance when contrasted with the baseline, can be interpreted as an increased interaction of the right with the LH. The longer timescale computations handled in the right STG could be fed to an internal model used by the left STG to better predict linguistic focus. At a more lexical level, the right dominance would not be as strong, longer timescale computations being less critical, which would explain the bilateral activity observed in T2.

Second, a dorsal stream was also observed. Bilateral activation was observed in the inferior parietal cortex, suggesting here again an interplay between the two hemispheres in this auditory-motor hub. The left dominant activation of the superior parietal lobule can be related with previous evidence for the role of this region in the processing of pointing (prosodic focus being a linguistic pointing act). Bilateral involvement of frontal regions was observed. The right dominance of the frontoopercular cortex region may reflect the processing of slow speech modulation. The involvement of the left inferior frontal gyrus is consistent with the view that this region is a parser of action structure. The bilateral involvement of the SMA further highlights the interaction with motor processes in focus detection.

Our results are therefore in favour of a dual route for the perception of prosodic focus, in which the observed interplay between the two hemispheres may reflect computations over different timescales. The longer timescale computations could provide predictions, which could facilitate shorter timescale computations. The observation of frontal activations could mean that motor simulation is necessary during prosodic focus detection.

Perceiving F_0 Variations Associated With Prosodic Focus: DCF_0R Parametric Analysis

An item-wise parametric analysis was performed to show the network of cerebral regions positively correlated with the normalized F_0 amplitude. The bilateral posterior part of the superior temporal gyrus was involved, as well as ventral regions including the left middle temporal gyrus, dorsal regions including the left inferior parietal gyrus, the bilateral inferior frontal gyrus, the bilateral insula, the bilateral superior frontal gyrus, the left premotor cortex, and the left cerebellum and the right thalamus.

Therefore, the parametric analysis shows the involvement of a dual network in the processing of F_0 amplitude,

similar to that involved in the detection of prosodic focus, although more left dominant. The DCF_0R parameter is related to pitch height detection, taking sentence declination into account, and could thus be one of the crucial acoustic cues to prosodic focus. The cerebral network correlated positively with this parameter, is more left lateralized than the cerebral network associated with prosodic focus detection (F vs. N). The left lateralization observed for the DCF_0R parameter could suggest that this parameter is more associated with a phonological distinction between focused and neutral utterances. As discussed earlier, the recruitment of the RH has been related to the processing of slowly changing acoustic cues (the AST hypothesis). These temporal cues are not present in the normalized F_0 amplitude value, which would confirm the specialization of the RH for the processing of long-term acoustic information. On the contrary, the acoustic cue lateralization hypothesis, which predicts that this parameter, associated with pitch but not with duration, should yield a RH asymmetry, is not supported by these results.

CONCLUSION

In this study, we have shown that the brain regions recruited in the detection of contrastive prosodic focus can be described as a network involving ventral and dorsal regions, slightly more right-dominant that the classical dual stream described in word or syllable perception by Hickok and Poeppel [2007]. The contrast between Focus and Neutral conditions revealed a network of ventral regions including the right superior temporal gyrus and the bilateral middle temporal gyrus and of dorsal regions including the bilateral inferior parietal gyrus, the left superior parietal gyrus, and the bilateral inferior frontal gyrus.

The finding of a difference in lateralization between the two (linguistic) prosodic conditions does not support the functional lateralization hypothesis which posits affective prosody to be handled by the RH and linguistic prosody by the LH. The involvement of the RH in the perisylvian cortex and more specifically in the posterior superior temporal gyrus could be explained by the processing and detection of slowly varying acoustic cues crucial in prosodic focus recognition. It is in line with an extended version of the AST hypothesis. Both conditions involved the processing of slowly varying acoustic cues, yet only the focus detection condition yielded a RH advantage. This last result can be interpreted in terms of increased activity and connectivity of the right temporal lobe with its left counterpart to check that the prosodic variations can be linguistically recognized as corrective focus. Longer timescale information processed in the right STG could potentially facilitate left STG computations over shorter temporal windows. The bilateral activation of the middle temporal gyrus can also be accounted for as the manifestation of an interplay between the two hemispheres with RH computations of slowly changing prosodic information helping to predict the lexical-level computa-

tions carried out in the left middle temporal gyrus. Similar coupled mechanisms in the left and right parietal and frontal lobes can be evoked to account for our findings.

When a parametric analysis is carried out using a parameter that captures pitch height and compensates for declination phenomena, the RH predominance is no longer present. This last result shows does not support the acoustic cue lateralization, which predicts a RH dominance for the processing of pitch. It is more consistent with the view that the DCF_0R parameter is associated with a phonological distinction between focused and neutral utterances.

Overall our results argue for a dual stream model of focus perception, which involves an interplay between the right and LHs, in which the computation of slowly changing prosodic cues in the RH dynamically feeds an internal model concurrently used by the LH to recognize prosodic focus. The observation of activations in frontal regions may reflect the use of feed-forward motor information during prosodic focus detection.

APPENDIX

Detailed corpus: sentences, phonetic transcription and translation into English.

REFERENCES

Andersson JL, Hutton C, Ashburner J, Turner R, Friston K (2001): Modeling geometric deformations in EPI time series. Neuroimage 13:903–919.

Astesano C, Besson M, Alter K (2004): Brain potentials during semantic and prosodic processing in French. Brain Res Cogn Brain Res 18:172–184.

Baum SR (1998): The role of fundamental frequency and duration in the perception of linguistic stress by individuals with brain damage. J Speech Lang Hear Res 41:31–40.

Belin P, Zilbovicius M, Crozier S, Thivard L, Fontaine A, Masure MC, Samson Y (1998): Lateralization of speech and auditory temporal processing. J Cogn Neurosci 10:536–540.

Boemio A, Fromm S, Braun A, Poeppel D (2005): Hierarchical and asymmetric temporal sensitivity in human auditory cortices. Nat Neurosci 8:389–395.

Bögels S, Schriefers H, Vonk W, Chwilla DJ (2011): Pitch accents in context: How listeners process accentuation in referential communication. Neuropsychologia 49:2022–2036.

Brådvik B, Dravins C, Holtas S, Rosen I, Ryding E, Ingvar DH (1991): Disturbances of speech prosody following right hemisphere infarcts. Acta Neurol Scand 84:114–126.

Buchel C, Holmes AP, Rees G, Friston KJ (1998): Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. Neuroimage 8:140–148.

Buchel C, Wise RJ, Mummery CJ, Poline JB, Friston KJ (1996): Nonlinear regression in parametric activation studies. Neuroimage 4:60–66.

Caplan D, Alpert N, Waters G, Olivieri A (2000): Activation of Broca's area by syntactic processing under conditions of concurrent articulation. Hum Brain Mapp 9:65–71.

Caplan D, Baker C, Dehaut F (1985): Syntactic determinants of sentence comprehension in aphasia. Cognition 21:117–175.

Caplan D, Hanna JE (1998): Sentence production by aphasic patients in a constrained task. Brain Lang 63:184–218.

Collina S, Marangolo P, Tabossi P (2001): The role of argument structure in the production of nouns and verbs. Neuropsychologia 39:1125–1137.

Delgutte B (1978): Technique for the perceptual investigation of F0 contours with application to French. J Acoust Soc Am 64:1319–1332.

Démonet JF, Thierry G, Cardebat D (2005): Renewal of the neurophysiology of language: Functional neuroimaging. Physiol Rev 85:49–95.

Di Cristo A (2000): Vers une modélisation de l'accentuation du français (seconde partie). J French Lang Studies 10:27–44.

Dohen M, Lœvenbruck H (2004): Pre-focal rephrasing, focal enhancement and postfocal deaccentuation in French. In 8th International Conference on Spoken Language Processing (ICSLP 04), île de Jeju (Corée), October 4–8, 2004, Vol. 1, 785–788.

Dohen M, Lœvenbruck H, Cathiard MA, Schwartz JL (2004): Visual perception of contrastive focus in reiterant French speech. Speech Commun 44:155–172.

Dohen M, Lœvenbruck H, Hill H (2009): Recognizing prosody from the lips: Is it possible to extract prosodic focus from lip features? In: Liew AW-C, Wang S, editors. Visual Speech Recognition: Lip Segmentation and Mapping. New York: Hershey. pp 416-438. ISBN 978-1-60566-186-5.

Fadiga L, Craighero L, D'Ausilio A (2009): Broca's area in language, action, and music, Ann NY Acad Sci 1169:448–458.

Fournier R, Gussenhoven C, Jensen O, Hagoort P (2010): Lateralization of tonal and intonational pitch processing: An MEG study. Brain Res 1328:79–88.

Friederici AD, Gorrell P (1998): Structural prominence and agrammatic theta-role assignment: A reconsideration of linear strategies. Brain Lang 65:253–275.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD, Turner R (1998): Event-related fMRI: Characterizing differential responses. Neuroimage 7:30–40.

Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RSJ (1995): Statistical parametric maps in functional imaging: A general linear approach. Hum Brain Mapp 2:189–210.

Friston KJ, Zarahn E, Josephs O, Henson RN, Dale AM (1999): Stochastic designs in event-related fMRI. Neuroimage 10:607–619.

Gandour J, Dzemidzic M, Wong D, Lowe M, Tong Y, Hsieh L, Satthamnuwong N, Lurito J (2003a): Temporal integration of speech prosody is shaped by language experience: An fMRI study. Brain Lang 84:318–336.

Gandour J, Tong Y, Talavage T, Wong D, Dzemidzic M, Xu Y, Li X, Lowe M (2007): Neural basis of first and second language processing of sentence-level linguistic prosody. Hum Brain Mapp 28:94–108.

Gandour J, Tong Y, Wong D, Talavage T, Dzemidzic M, Xu Y, Li X, Lowe M (2004): Hemispheric roles in the perception of speech prosody. Neuroimage 23:344–357.

Gandour J, Wong D, Dzemidzic M, Lowe M, Tong Y, Li X (2003b): A cross-linguistic fMRI study of perception of intonation and emotion in Chinese. Hum Brain Mapp 18:149–157.

Geiser E, Zaehle T, Jancke L, Meyer M (2008): The neural correlate of speech rhythm as evidenced by metrical speech processing. J Cogn Neurosci 20:541–552.

Giraud A, Kleinschmidt A, Poeppel D, Lund T, Frackowiak R, Laufs H (2007): Endogenous cortical rhythms determine cerebral specialization for speech perception and production. Neuron 56:1127–1134.

Gussenhoven C (2007): Types of focus in English. Topic and Focus, Springer. pp83–100.

Hesling I, Dilharreguy B, Clement S, Bordessoules M, Allard M (2005): Cerebral mechanisms of prosodic sensory integration using low-frequency bands of connected speech. Hum Brain Mapp 26:157–169.

Hickok G, Poeppel D (2000): Towards a functional neuroanatomy of speech perception. Trends Cogn Sci 4:131–138.

Hickok G, Poeppel D (2004): Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. Cognition 92:67–99.

Hickok G, Poeppel D (2007): The cortical organization of speech processing. Nat Rev Neurosci 8:393–402.

Hickok G, Houde J, Rong F (2011): Sensorimotor integration in speech processing: Computational basis and neural organization. Neuron 69:407–422.

Hruska C, Alter K, Steinhauer K, Steube A (2001): Misleading dialogs: Human's brain reaction to prosodic information. In: Cave C, Guaitella I, Santi S, editor. Orality and gestures. Interactions et comportements multimodaux dans la communication. Paris: L'Hartmattan.

Johnson SM, Clifton C, Breen M, Morris J (2003): An ERP investigation of prosodic and semantic focus. Poster presented at Cognitive Neuroscience, New York City.

Just MA, Carpenter PA, Keller TA, Eddy WF, Thulborn KR (1996): Brain activation modulated by sentence comprehension. Science 274:114–116.

Kotz SA, Schwartze M (2010): Cortical speech processing unplugged: A timely subcortico-cortical framework. Trends Cogn Sci 14:392–399.

Ladd DR (1980): The structure of intonational meaning: Evidence from English. Bloomington: Indiana University Press.

Ladd DR (1996): Intonational phonology. Vol. 79. New York: Cambridge University Press.

Lau E, Phillips C, Poeppel D (2008): A cortical network for semantics:(de) constructing the N400. Nat Rev Neurosci 9:920–933.

Liberman M, Pierrehumbert J (1984): Intonational invariance under changes in pitch range and length: Language Sound Structure: Studies in Phonology Presented to Morris Halle by His Teacher and Students. The MIT Press.

Lœvenbruck H, Baciu M, Segebarth C, Abry C (2005): The left inferior frontal gyrus under focus: An fMRI study of the production of deixis via syntactic extraction and prosodic focus. J Neurolinguist 18:237–258.

Lœvenbruck H, Dohen M, Vilain C (2009): Pointing is 'special.' In: Fuchs S, Lœvenbruck H, Pape D, Perrier P, editors. Some Aspects of Speech and the Brain. Peter Lang. pp 211–258. ISBN 978-3-631-57630-4.

Luks TL, Nusbaum HC, Levy J (1998): Hemispheric involvement in the perception of syntactic prosody is dynamically dependent on task demands. Brain Lang 65:313–332.

Magne C, Astésano C, Lacheret-Dujour A, Morel M, Alter K, Besson M (2005): On-line processing of "pop-out" words in spoken French dialogues. J Cogn Neurosci 17:740–756.

Meyer M, Alter K, Friederici A (2003): Functional MR imaging exposes differential brain responses to syntax and prosody during auditory sentence comprehension. J Neurolinguist 16:277–300.

Meyer M, Alter K, Friederici AD, Lohmann G, von Cramon DY (2002): FMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. Hum Brain Mapp 17:73–88.

Meyer M, Steinhauer K, Alter K, Friederici AD, von Cramon DY (2004): Brain activity varies with modulation of dynamic pitch variance in sentence melody. Brain Lang 89:277–289.

Meyer M, Zaehle T, Gountouna VE, Barron A, Jancke L, Turk A (2005): Spectro-temporal processing during speech perception involves left posterior auditory cortex. NeuroReport 16:1985.

Oldfield RC (1971): The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9:97–113.

Overath T, Kumar S, von Kriegstein K, Griffiths T (2008): Encoding of spectral correlation over time in auditory cortex. J Neurosci 28:13268–13273.

Pell MD (1998): Recognition of prosody following unilateral brain lesion: Influence of functional and structural attributes of prosodic contours. Neuropsychologia 36:701–715.

Pell MD (1999): The temporal organization of affective and non-affective speech in patients with right-hemisphere infarcts. Cortex 35:455–477.

Pell MD, Baum SR (1997): Unilateral brain damage, prosodic comprehension deficits, and the acoustic cues to prosody. Brain Lang 57:195–214.

Perkins JM, Baran JA, Gandour J (1996): Hemispheric specialization in processing intonation contours. Aphasiology 10:343–362.

Pickering M, Garrod S (2007): Do people use language production to make predictions during comprehension? Trends Cogn Sci 11:105–110.

Poeppel D (2003): The analysis of speech in different temporal integration windows: Cerebral lateralization as 'asymmetric sampling in time'. Speech Commun 41:245–255.

Poeppel D, Idsardi W, Van Wassenhove V (2008): Speech perception at the interface of neurobiology and linguistics. Philos Trans R Soc B: Biol Sci 363:1071–1086.

Rauschecker J (2011): An expanded role for the dorsal auditory pathway in sensorimotor control and integration. Hearing Res 271:16–25.

Rauschecker J, Scott S (2009): Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. Nat Neurosci 12:718–724.

Rauschecker J, Tian B (2000): Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci USA 97:11800–11806.

Rigalleau F, Baudiffier V, Caplan D (2004): Comprehension of sentences with stylistic inversion by French aphasic patients. Brain Lang 89:142–156.

Rizzolatti G, Arbib M (1998): Language within our grasp. Trends Neurosci 21:188–194.

Sato M, Tremblay P, Gracco VL (2009): A mediating role of the premotor cortex in phoneme segmentation. Brain Lang 111:1–7.

Schirmer A, Alter K, Kotz SA, Friederici AD (2001): Lateralization of prosody during language production: A lesion study. Brain Lang 76:1–17.

Schwartz MF, Linebarger MC, Saffran EM, Pate DS (1987): Syntactic transparency and sentence interpretation in aphasia. Lang Cogn Process 2:85–113.

Schwartz J-L, Ménard L, Basirat A, Sato M (in press): The perception-for-action-control theory (PACT): A perceptuo-motor theory of speech perception. J Neurolinguistics.

Scott SK, Johnsrude IS (2003): The neuroanatomical and functional organization of speech perception. Trends Neurosci 26:100–107.

Selkirk EO (1978): On prosodic structure and its relation to syntactic structure. Paper presented at the Paper presented at the Nordic Prosody II, Triondheim.

Selkirk EO (2008): Contrastive focus, givenness and the unmarked status of "discourse-new". Acta Linguistica Hungarica 55:331–346.

Shipley-Brown F, Dingwall WO, Berlin CI, Yeni-Komshian G, Gordon-Salant S (1988): Hemispheric processing of affective and linguistic intonation contours in normal subjects. Brain Lang 33:16–26.

Shtyrov Y, Pihko E, Pulvermuller F (2005): Determinants of dominance: Is language laterality explained by physical or linguistic features of speech? Neuroimage 27:37–47.

Skipper JI, Nusbaum HC, Small SL (2005): Listening to talking faces: Motor cortical activation during speech perception. Neuroimage 25:76–89.

Stiller D, Gaschler-Markefski B, Baumgart F, Schindler F, Tempelmann C, Heinze HJ, Scheich H (1997): Lateralized processing of speech prosodies in the temporal cortex: A 3-T functional magnetic resonance imaging study. Magma 5:275–284.

Toepel U, Pannekamp A, Alter K (2007): Catching the news: Processing strategies in listening to dialogs as measured by ERPs. Behav Brain Funct 3:53.

Tong Y, Gandour J, Talavage T, Wong D, Dzemidzic M, Xu Y, Li X, Lowe M (2005): Neural circuitry underlying sentence-level linguistic prosody. Neuroimage 28:417–428.

Twist DJ, Squires NK, Spielholz NI, Silverglide R (1991): Event-related potentials in disorders of prosodic and semantic linguistic processing. Cogn Behav Neurol 4:281.

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002): Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. NeuroImage 15:273–289.

Van Lancker D, Sidtis JJ (1992): The identification of affective-prosodic stimuli by left- and right-hemisphere-damaged subjects: All errors are not created equal. J Speech Hear Res 35:963–970.

Vigneau M, Beaucousin V, Hervé P-Y, Jobard G, Petit L, Crivello F, Mellet E, Zago L, Mazoyer B, Tzourio-Mazoyer N (2011): What is right-hemisphere contribution to phonological, lexico-semantic, and sentence processing?: Insights from a meta-analysis. NeuroImage 54:577–593.

Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O, Mazoyer B, Tzourio-Mazoyer N (2006): Meta-analyzing left hemisphere language areas: Phonology, semantics, and sentence processing. NeuroImage 30:1414–1432.

Von Kriegstein K, Smith DRR, Patterson RD, Kiebel SJ, Griffiths TD (2010): How the human brain recognizes speech in the context of changing speakers. J Neurosci 30:629–638.

Webster J, Franklin S, Howard D (2001): An investigation of the interaction between thematic and phrasal structure in nonfluent agrammatic subjects. Brain Lang 78:197–211.

Weintraub S, Mesulam MM, Kramer L (1981): Disturbances in prosody. A right-hemisphere contribution to language. Arch Neurol 38:742–744.

Welby PS (2003): The slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation. The Ohio State University.

Whitworth A (1995): Characterising thematic role assignment in aphasic sentence production: Procedures for elicited and spontaneous output. Inter J Lang Commun Disorders 30:384–399.

Wiethoff S, Wildgruber D, Kreifelts B, Becker H, Herbert C, Grodd W, Ethofer T (2008): Cerebral processing of emotional

prosody—Influence of acoustic parameters and arousal. Neuro-Image 39:885–893.

Wildgruber D, Ackermann H, Kreifelts B, Ethofer T (2006): Cerebral processing of linguistic and emotional prosody: fMRI studies. Prog Brain Res 156:249–268.

Wildgruber D, Hertrich I, Riecker A, Erb M, Anders S, Grodd W, Ackermann H (2004): Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation. Cereb Cortex 14:1384–1389.

Xu Y, Gandour J, Talavage T, Wong D, Dzemidzic M, Tong Y, Li X, Lowe M (2006): Activation of the left planum temporale in pitch processing is shaped by language experience. Hum Brain Mapp 27:173–183.

Zatorre RJ, Belin P, Penhune VB (2002): Structure and function of auditory cortex: Music and speech. Trends Cogn Sci 6: 37–46.

Zatorre RJ, Evans AC, Meyer E, Gjedde A (1992): Lateralization of phonetic and pitch discrimination in speech processing. Science 256:846–849.

Zatorre RJ, Gandour JT (2008): Neural specializations for speech and pitch: Moving beyond the dichotomies. Philos Trans R Soc B: Biol Sci 363:1087.