



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SciVerse ScienceDirect

Journal homepage: [www.elsevier.com/locate/cortex](http://www.elsevier.com/locate/cortex)

## Research report

## Regional heterogeneity in the processing and the production of speech in the human planum temporale

Pascale Tremblay<sup>a,\*,1</sup>, Isabelle Deschamps<sup>b,c</sup> and Vincent L. Gracco<sup>b,c,d</sup><sup>a</sup> Center for Mind & Brain Sciences (CIMeC), The University of Trento, Italy<sup>b</sup> McGill University, Faculty of Medicine, School of Communication Sciences and Disorders, Montreal, Canada<sup>c</sup> Centre for Research on Brain, Language & Music, McGill University, Canada<sup>d</sup> Haskins Laboratories, New Haven, CT, USA

## ARTICLE INFO

## Article history:

Received 16 December 2010

Reviewed 9 May 2007

Revised 12 June 2011

Accepted 31 August 2011

Action editor Stefano Cappa

Published online 29 September 2011

## Keywords:

Posterior temporal plane

Word production

Word repetition

Whispering

Auditory feedback

## ABSTRACT

**Introduction:** The role of the left planum temporale (PT) in auditory language processing has been a central theme in cognitive neuroscience since the first descriptions of its leftward neuroanatomical asymmetry. While it is clear that PT contributes to auditory language processing there is still some uncertainty about its role in spoken language production.

**Methods:** Here we examine activation patterns of the PT for speech production, speech perception and single word reading to address potential hemispheric and regional functional specialization in the human PT. To this aim, we manually segmented the left and right PT in three non-overlapping regions (medial, lateral and caudal PT) and examined, in two complementary experiments, the contribution of exogenous and endogenous auditory input on PT activation under different speech processing and production conditions.

**Results:** Our results demonstrate that different speech tasks are associated with different regional functional activation patterns of the medial, lateral and caudal PT. These patterns are similar across hemispheres, suggesting bilateral processing of the auditory signal for speech at the level of PT.

**Conclusions:** Results of the present studies stress the importance of considering the anatomical complexity of the PT in interpreting fMRI data.

© 2011 Elsevier Ltd. All rights reserved.

## 1. Introduction

The planum temporale (PT) is a large sheet-like cortical area with a roughly triangular shape located on the superior temporal gyrus (STG), posterior to the primary auditory area (PAC) which is located on Heschl's sulcus (von Economo and

Horn, 1930; Pfeifer, 1936; Galaburda and Sanides, 1980). The contribution of the left PT to auditory language processing has been a central theme in cognitive neuroscience since the first descriptions of its leftward asymmetry (Geschwind and Levitsky, 1968; Galaburda et al., 1978; Steinmetz and Galaburda, 1991; for a systematic review, see Shapleske

\* Corresponding author. Center for Mind & Brain Sciences (CIMeC), Università Degli Studi di Trento, Via delle Regole, 101, I – 38100 Mattarello (TN), Italy.

E-mail address: [Pascale.Tremblay@fmed.ulaval.ca](mailto:Pascale.Tremblay@fmed.ulaval.ca) (P. Tremblay).

<sup>1</sup> Present address: Faculté de Médecine, Département de réadaptation, Université Laval Pavillon Ferdinand-Vandry, Bureau 4462, 1050 avenue de la Médecine, Québec (QC), Canada G1V 0A6.

0010-9452/\$ – see front matter © 2011 Elsevier Ltd. All rights reserved.

doi:10.1016/j.cortex.2011.09.004

et al., 1999). It has been hypothesized that this asymmetry reflects a functional specialization of the left cerebral hemisphere for language (Goulven et al., 2003; Dorsaint-Pierre et al., 2006; but see also Marshall, 2000). Additional support for this hypothesis comes from lesions studies that have shown that damage in or around PT is associated with auditory language comprehension deficits (Tanaka et al., 1987; Praamstra et al., 1991; Caplan et al., 1995). Furthermore, Jacquemot et al. (2003) have shown that the left STG/lateral PT selectively responds to language-specific phonological manipulations. In other neuroimaging experiments PT activation has been observed during passive listening to speech sounds (Petersen et al., 1988; Wise et al., 1991, 2001; Zatorre et al., 1992; Mazoyer et al., 1993; Binder et al., 1996, 1997, 2000; Friederici et al., 2000; Belin et al., 2002; Callan et al., 2006; Wilson et al., 2008) with increased activation in background noise (Wang et al., 2008). However, PT also responds to non-language perceptual processing including melody (Griffiths et al., 1998; Zatorre et al., 1994), tones (Specht and Reul, 2003; Hugdahl et al., 1999; Jancke et al., 2003; Binder et al., 1996), musical instruments (Hugdahl et al., 1999) and masking noise (Wise et al., 2001). Unilateral and/or bilateral STG lesions that include PT result in impaired recognition of a wide range of non-verbal environmental sounds (including human non-verbal sounds such as crying), manmade and natural non-living sounds and animal sounds (Schnider et al., 1994). In terms of a possible functional specialization, passive and active listening results in greater PT activation bilaterally for tones compared to words, suggesting a general auditory processing function rather than language-specific processing (Binder et al., 1996, 1997). Consistent with these and other more recent findings (e.g., Warren et al., 2005a, 2005b; Kumar et al., 2007) it has been suggested that the bilateral PT acts as a computational hub whose function is to disambiguate complex sounds by isolating different properties of the acoustic objects (e.g., temporal and spectral information) and matching them to stored templates (Griffiths and Warren, 2002). Despite some findings incompatible with the computational hub hypothesis (see for example Hall and Plack, 2009), it is clear that PT is involved in the processing of sounds, including speech, a function consistent with its connectivity pattern and anatomical location.

In addition to serving as a computational hub for auditory processing, PT may be involved in sensorimotor (auditory–motor) transformation. According to this view, sensorimotor transformation is necessary to establish parity between motor and perceptual units for speech (as in the Motor Theory of Speech Perception; Liberman and Mattingly, 1985). Parity functions to link motor (articulatory) representations to their auditory consequences. Sensorimotor transformation is necessary for both speech perception (as in the Motor Theory of Speech Perception; Liberman and Mattingly, 1985) and speech production (Warren et al., 2005a, 2005b; Hickok and Poeppel, 2007). This hypothetical function is most likely accomplished through connections to and from motor regions through the arcuate fasciculus (AF) (i.e., the dorsal stream network) (Hickok et al., 2000, 2003; Buchsbaum et al., 2001; Warren et al., 2005a, 2005b; Saito et al., 2006; see also Hickok and Poeppel, 2000, 2004, 2007; Schmahmann and Pandya, 2006; Schmahmann et al., 2007; Saur et al., 2008) as well as other networks that have yet to be

identified. Support for the role for PT in sensorimotor transformation comes mainly from imaging studies showing PT activation during overt speech production (Petersen et al., 1988; Wise et al., 1991; Karbe et al., 1998; Schulz et al., 2005; Saito et al., 2006; Tourville et al., 2008; Bohland and Guenther, 2006; Riecker et al., 2008; Zheng et al., 2010; Peschke et al., 2009; Dhanjal et al., 2008), but also during silent speech production or speech rehearsal, which does not involve self-generated auditory feedback (Hickok et al., 2000, 2003; Buchsbaum et al., 2001; Wise et al., 2001; Huang et al., 2001; Papathanassiou et al., 2000; Okada et al., 2003; Shergill et al., 2002; Callan et al., 2006; Pa and Hickok, 2008). Furthermore, it has been shown that an area of the posterior temporal–parietal junction, often referred to as area Spt (which includes the caudal part of PT) responds more strongly to sub-vocal rehearsal of auditory stimuli than to the perception of auditory stimuli (e.g., Buchsbaum et al., 2001; Hickok et al., 2003; Hickok et al., 2009). Consistent with these results, left PT activation co-varies with changes in the left motor cortex activation during whispered speech (Paus et al., 1996), providing indirect evidence of a connection between motor regions and PT, and support for a role for PT in sensorimotor transformation for speech.

While brain regions can support more than one function, it is possible that the apparent heterogeneity of functions supported by PT is related to the fact that the human PT, like the non-human primate PT, contains several cortical fields (e.g., von Economo and Horn, 1930; Galaburda and Sanides, 1980; Rivier and Clarke, 1997; Tardif and Clarke, 2001; Scheich et al., 1998; Sweet et al., 2005; Fullerton and Pandya, 2007). Just posterior to the transverse temporal sulcus, in a region corresponding to the rostral PT, two regions have been identified by Rivier and Clarke (1997), the lateral area (LA) and the posterior (and more medial) area (PA). Areas LA and PA are both part of the internal and external parabelt areas identified for example by Sweet et al. (2005) in humans. The caudal most region of PT, often referred to as the temporoparietal area (Tpt), and sometimes as the vertical planum temporale (vPT)<sup>2</sup>, is located outside of the auditory parabelt (Sweet et al., 2005). Since different cortical fields are likely to support different functions, it seems reasonable to hypothesize that PT is a functionally heterogeneous region, a claim that is supported by the literature. In sum, PT appears to be involved in auditory speech processing and in transforming auditory signals into speech motor representations. In addition, PT may also be activated as a result of auditory re-afference to monitor feedback during speech production. In order to understand the role of PT in speech processing, the contribution of each of these three sources of activation needs to be evaluated.

In the present study, we evaluated the hypothesis that the anatomically distinct regions of the left and right PT are differentially activated depending on the source and function of the auditory processing involved. We examined, in two complementary studies, the contribution of exogenous and endogenous auditory input on PT activation under different speech processing and production conditions. Our main objective was to investigate whether different regions of the human PT

<sup>2</sup> It should be noted that Tpt and vPT are not identical: the human Tpt extends onto the parietal convexity (Galaburda et al., 1978; Spocter et al., 2010), while the vPT does not.

contribute to the perception and production of speech in distinct ways. In the first study, we examined the processing of internally generated auditory stimulation by comparing whispered speech production with and without feedback. In the second study, we compared the processing of internally and externally generated auditory stimulation by comparing auditory word repetition (AWR) and word reading aloud. In general PT activation appears to be functionally differentiated suggesting unique contributions to speech perception and production.

## 2. Experiment 1

### 2.1. Material and methods

#### 2.1.1. Participants

Ten healthy adults (mean  $25.1 \pm 3.8$  years) balanced for gender comprised the experimental group. All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) and were native speakers of Canadian English. All participants had normal or corrected-to-normal vision, and reported no history of speech, language or learning difficulties. Participants were screened for any contraindication to magnetic resonance imaging (MRI). Informed written consent was obtained from each participant. The study was approved by the Magnetic Resonance Research Committee (MRRC) and the Montreal Neurological Institute (MNI) Research Ethics Committee.

#### 2.1.2. Experimental procedures

The experiment consisted of two active speech conditions that were presented in pseudo-random blocks of two or three trials. Subjects were instructed to whisper a printed word presented on the back-projected screen under either normal feedback (Whisper) or in the presence of masking noise (Whisper masked). The masking noise level was adjusted for each participant such that it was impossible for him or her to hear his or her own voice while the masking was present. The noise began at the time participants were instructed to start whispering. The rationale for using whispered speech as a proxy of natural speech production, is that (1) in whispered speech subjects articulate in a near normal manner, and (2) the self-generated auditory feedback can be completely masked with noise, because of the absence of bone-conducted feedback. These two speaking conditions involved similar motor processes but differed in terms of the presence of auditory feedback. Importantly, participants were trained not to change their articulation in the masking condition, to avoid a Lombard effect, characterized by an increase in voice intensity in the presence of background noise (Lombard, 1911, as cited in H. Lane and B. Tranel, 1971). During the practice session, participants were asked to produce visually presented words in the presence of noise and no noise and asked to keep articulation and voice level constant. However, it should be noted that the acoustic recordings were not of high enough quality to allow for acoustical analyses, which would have demonstrated whether or not there was a Lombard effect. We are nevertheless confident that subjects maintained their voice constant between the two conditions because we did not observe differences in motor cortex

activation between the tasks. Increased speaking intensity would typically translate into more effort and lead to increase motor activation – this was not observed. The lack of increased activation in the primary motor area in the Whisper Masked condition is shown in [Supplementary Fig. S2](#).

The printed stimuli used for the speech condition were three-syllable nouns presented on a screen one at a time (see [Supplemental material](#) for the list of all stimuli). Each experimental trial began with the presentation of the visual instruction (ex. Whisper “tomato”) presented during the volume acquisition (3.3 sec) using SuperLabPro (Cedrus Corp, CA, USA). The visual instruction remained on the screen for the duration of the volume acquisition in order to limit the need for participants to sub-vocally rehearse the words. Participants were instructed to wait until the stimulus disappeared to respond, which occurred at the end of the volume acquisition. All responses occurred during a silent interval of 6.2 sec – a delay that was chosen such that at the time of the consecutive volume acquisition, the hemodynamic response to speech production would be maximal (that is, 5–6 sec after onset of speech production).

Fifty trials of the Whisper condition, and 50 trials of the Whisper Masked conditions were acquired in total. The baseline condition, which was interleaved with the experimental conditions, was a visual fixation condition during which participants fixated on the word “rest” (30 trials). This was included for two reasons: (1) to control for the visual stimuli that occurred during the experimental conditions, (2) to allow for an examination of the effects of each task against a non-speech task, as opposed to looking at contrasts only. A passive listening to masking noise condition (MN) was also added to the experimental paradigm (30 trials). During this task, participants fixated on the word “rest”. The noise was presented during the delay in TR only. During the scanning session, participants also produced additional tasks that were not analyzed for the current study (tongue raising and lowering, jaw openings and closings, silent articulation).

#### 2.1.3. Image acquisition

The data were acquired on a 1.5 T Siemens Sonata MRI scanner at the MNI (Montreal, Canada). In order to eliminate movement artifacts associated with producing speech in the scanner, a sparse image acquisition technique (Eden et al., 1999; Edmister et al., 1999; Gracco et al., 2005) was used. A silent period (6.2 sec) was interleaved between each volume acquisition. All responses occurred during the silent period. The subjects wore MR compatible headphones (Commander XG, Resonance technology, CA, USA) and their responses were recorded through a MR compatible microphone attached to the headphones and digitized directly onto a Toshiba laptop computer. Thirty-nine axial slices (whole brain coverage) oriented parallel to the AC–PC line (thickness = 4 mm, no gap, FOV =  $256 \times 256$  mm<sup>2</sup>, matrix =  $64 \times 64$ ) were acquired in 3.3 sec using a multi-slice echo-planar imaging (EPI) sequence (TE = 50 msec, TR = 9.5 sec, delay in TR = 6.2 sec). The delay in TR occurred following each volume acquisition. The slices had a spatial resolution of  $4 \times 4 \times 4$  mm. Two experimental runs (13 min each) resulted in the acquisition of 266 T2\*-weighted BOLD images acquired in a descending order. High-resolution T1-weighted volumes were acquired for

anatomical localization (matrix  $256 \times 256$  mm, 176 slices,  $1 \times 1 \times 1$  mm, no gap, TE = 9.2 msec, TR = 22 msec). The participants' head was immobilized by means of a vacuum-bag filled with polystyrene balls, which was fitted around the subject's head.

#### 2.1.4. Image analysis

Images were spatially registered, motion-corrected, despiked, mean-normalized and smoothed with a Gaussian 6-mm FWHM filter using AFNI (Cox, 1996). There were separate regressors for the experimental conditions (Whisper, Whisper masked), as well as the control condition (Listen to masking noise). Additional regressors were the mean, linear, and quadratic trend components, the six motion parameters ( $x$ ,  $y$ ,  $z$  and roll, pitch and yaw). A linear least squares model was used to establish a fit to each time point of the HRF for each condition. For the second level (group) analysis, subject data were transformed into stereotaxic space using the MNI305 template (Collins et al., 1994), and re-sampled to  $3 \times 3 \times 3$  mm. The group analyses were performed on the subjects' beta values resulting from the first level analysis. We conducted a two-way mixed model ANOVA with repeated measurements on the task (Whisper, Whisper Masked). Subjects were entered in the analysis as a random factor. A cluster correction for multiple comparisons across the brain was implemented in the AFNI program 3dClusterSim. Based on 3dClusterSim results, we determined that a family-wise error (FWE) rate of  $p < .05$  is achieved with a minimum cluster size of 44 contiguous voxels each significant at  $p < .01$ .

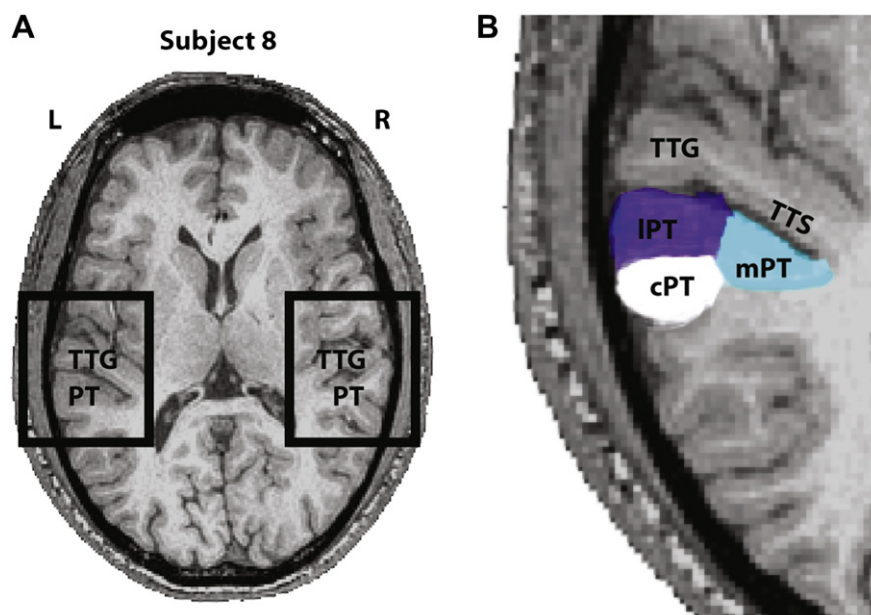
#### 2.1.5. Region of interest (ROI) analysis

Three anatomically defined regions of interest were created to parcellate the left and the right PT. Anatomical definitions were determined with reference to Rivier and Clarke (1997). For each subject, PT was manually segmented into three non-

overlapping regions (medial, lateral and caudal) in the following manner: PT was first split into two regions (rostral and caudal) with the transverse temporal sulcus as the rostral boundary. If two transverse temporal gyri were present, we used the gray matter caudal to the second gyri. The rostral region was then split into two halves (lateral and medial), roughly corresponding to Rivier and Clarke (1997) areas LA and PA respectively. The caudal boundary of the caudal PT was the ascending ramus of the Sylvian fissure. An example of a parcellation is shown on Fig. 1(A, B). For each subject, we calculated the mean percentage of signal change for each condition and PT region, as well as difference scores between the two whispering conditions (Whisper–Whisper Masked) and entered these values into a 2-way ANOVA with repeated measurements, with factor ROI (Medial, Lateral, Caudal PT) and Hemisphere (left, right). FDR corrected ( $i = 3$ ;  $q = .05$ ) post hoc tests (paired sample  $t$ -tests) were used to further examine the effect of ROI.

#### 2.1.6. SNR analyses

To evaluate the sensitivity of our experimental design, following Parrish et al. (2000) and Dick et al. (2009) we conducted a series of simulations that determined, for various SNR values, the probability of identifying a signal change of magnitude of .5% and 1%. For each participant, we then computed a signal-to-noise ratio (SNR) map of PT using AFNI. SNR was defined here as the ratio of the signal's mean to the standard deviation (SD) of the signal. The simulation was conducted using R. We derived a synthetic model of the expected activity by convolving the boxcar design with a canonical HRF (mean = 1). To this, varying amounts of smoothed noise (modeled by an autoregressive moving average) were added to obtain different SNR ratios. This process was repeated for 10,000 iterations for each noise level, where each iteration assigned a different randomly generated noise value. We then determined the probability of finding a reliable



**Fig. 1** – Example of a representative PT parcellation. **A.** Axial view of subject 8 brain showing the superior temporal plane. **B.** Zoomed view on the left superior temporal area showing the parcellation in PT. Legend: mPT = medial PT, IPT = lateral PT, cPT = caudal PT, TTG = transverse temporal gyrus, TTS = transverse temporal sulcus.

correlation for that noise level by evaluating the “boxcar-predictor-plus-noise” data against the original boxcar predictor via regression (see Parrish et al., 2000 for details of method). This simulation method indicates the power to reliably detect a given MR percent signal change as a function of SNR. Such a simulation determines whether the design has sufficient power to detect results in cases where null results are reported.

## 2.2. Results

### 2.2.1. Behavioral data

No errors were committed in the whispering tasks.

### 2.2.3. Neuroimaging results

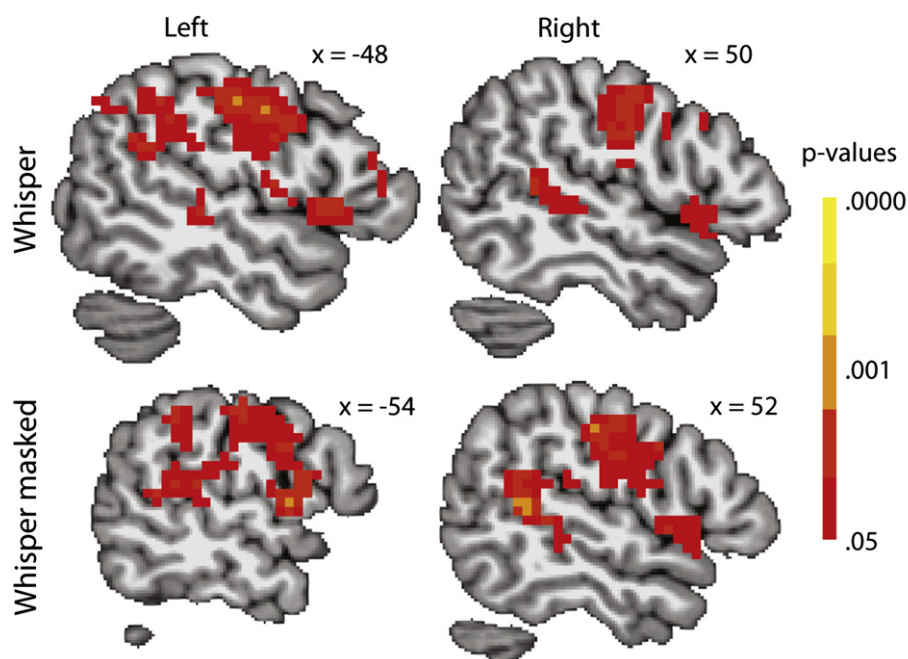
**2.2.2.1. SNR ANALYSES.** The mean SNR in the bilateral PT for the group was 69.8 ( $\pm 9.7$  SD). The power simulations determined that, given a power of .80, the minimum SNR to detect a MR signal change of .5% was 56, whereas the minimum SNR to detect a MR signal change of 1% was 27 (Supplementary Fig. S1). These results show that the actual SNR of the study was sufficient to detect signal changes in PT as low as .5%.

**2.2.2.2. Voxel-based analyses.** As illustrated in Fig. 2, the contrast of each whispering condition with the resting baseline revealed activation in the ventral precentral gyrus (including PMv and M1) bilaterally, left IFG pars opercularis (IFGop), left insula, transverse temporal gyrus and PT bilaterally, and cerebellum, lobules V and VI, bilaterally. In order to identify regions active for Whisper Masked above and beyond listening to masking noise alone, we also subtracted the Noise condition from Whisper Masked. This contrast revealed

activation in the ventral precentral gyrus (including PMv and M1) bilaterally, left IFGop. The contrast of the masking noise condition with the resting baseline revealed strong activation in transverse temporal gyrus and PT bilaterally. A list of all areas of activation, for each condition, is provided in Table 1.

**2.2.2.3. ROIs.** First, we computed a difference score between the two whispering conditions (Whisper–Whisper Masked), for each ROI and each hemisphere. These scores were submitted to a 2-way ANOVA with repeated measurements, with factor ROI (lateral PT, medial PT, caudal PT) and Hemisphere (left, right). Results show a main effect of Region [ $F_{(2,9)} = 5.91$ ,  $p = .011$ ], and a main effect of Hemisphere [ $F_{(1,9)} = 19.56$ ,  $p = .002$ ], but no interaction [ $F_{(2,18)} = .14$ ,  $p = .88$ ]. FDR-corrected paired sample t-tests indicate that the medial PT was significantly different from the lateral PT [ $t_{(9df)} = 3.11$ ,  $p = .016$ ]. While the lateral PT was significantly more active for Whisper than Whisper masked, the opposite effect was found at the level of the medial PT. There was no difference between the two whisper conditions at the level of the caudal PT. These results are illustrated in Fig. 3. The laterality effect was due to an overall stronger difference between the Whisper conditions in the left PT compared to the right PT.

We also computed the difference between the Whisper Masked condition and the Noise condition (Whisper Masked–Noise listening). These difference scores were submitted to a 2-way ANOVA with repeated measurements, with factor ROI (lateral PT, medial PT, caudal PT) and Hemisphere (left, right). Results show a marginally significant laterality effect, with an overall stronger difference (Whisper Masked > Noise listening) in the left than in the right PT [ $F_{(1,9)} = 5.03$ ,  $p = .052$ ].



**Fig. 2 – Cluster-corrected group-level activation (FWE rate of  $p < .05$ , achieved with a minimum cluster size of 44 contiguous voxels each significant at  $p < .01$ ) from Experiment 1 for the contrast of each Whispering condition against a resting baseline. Data are overlaid on axial views of a participant's T1-weighted MRI transformed into stereotaxic space using the MNI305 template. All coordinates are in MNI space.**

**Table 1 – Table of statistically significant activation at the whole-brain level, corrected for multiple comparison (FWE rate of  $p < .05$ , achieved with a minimum cluster size of 44 contiguous voxels each significant at  $p < .01$ ), for Experiment 1, for each condition compared against a resting baseline: (1) Whisper, (2) Whisper Masked, (3) Listening to masking noise, and (4) for the conjunction of the two whispering condition.**

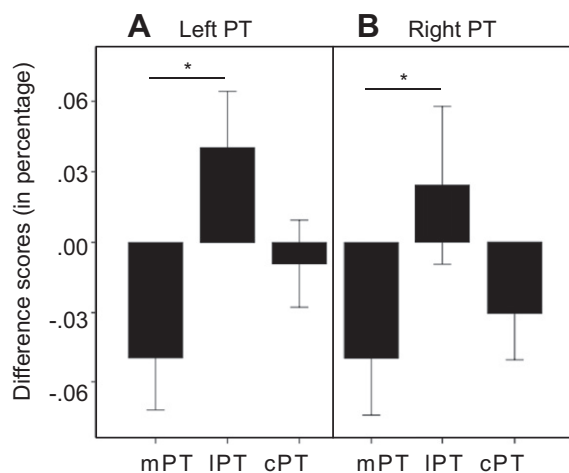
Description	Hemi	x	y	z	Voxels	p value
<b>1. Whisper</b>						
Precentral gyrus and sulcus, postcentral gyrus (including PMv and M1/S1), extending caudally into the supramarginal gyrus, and rostrally into the inferior frontal gyrus, pars opercularis and medially into the insula.	L	–50	–17	33	156	.00008
Precentral gyrus and central sulcus (M1/S1).	R	44	–11	42	104	.00017
Inferior frontal gyrus, pars opercularis and triangularis, and transverse temporal gyrus.	L	–53	2	12	80	.00256
Posterior cingulate gyrus.	R	2	–8	27	76	.00000
Anterior cingulate, medial frontal gyrus (pre-SMA).	L	–5	–8	63	48	.00005
<b>2. Whisper with masked feedback</b>						
Medial occipital lobe including calcarine fissure.	R	17	–83	–1	492	.00000
Precentral gyrus and sulcus, and postcentral gyrus (including PMv and M1/S1), extending rostrally into the inferior frontal gyrus, pars opercularis.	L	–50	–23	39	137	.00004
Supplementary motor area.	L	–5	–5	60	117	.00002
Precentral gyrus and sulcus, and postcentral gyrus (including PMv and M1/S1), extending rostrally into the inferior frontal gyrus, pars opercularis.	R	59	2	18	90	.00002
Cerebellum, lobules V and VI.	R	17	–59	–19	57	.00001
Inferior frontal gyrus, pars opercularis.	L	–56	–2	9	51	.00016
Superior temporal sulcus and gyrus, caudal PT.	R	59	–41	12	46	.00004
<b>3. Whisper masked minus listening to masking noise</b>						
Precentral gyrus and sulcus, and postcentral gyrus (including PMv and M1/S1), extending caudally into the supramarginal gyrus, rostrally into the inferior frontal gyrus, pars opercularis.	L	–53	–8	39	74	.000299577
<b>4. Listen to masking noise</b>						
STG, including the planum temporale and transverse temporal gyrus.	R	50	–26	18	67	.00030
STG, including the planum temporale and transverse temporal gyrus.	L	–38	–32	12	53	.00071
Frontal operculum, anterior insula.	L	–29	23	15	48	.00003
Calcarine fissure, occipito-temporal sulcus, lingual gyrus.	L	–14	–65	–1	45	.00003
Left and right precuneus.	R	2	–47	54	45	.00018
<b>5. Whisper <math>\cap</math> whisper masked</b>						
Precentral gyrus and sulcus, postcentral gyrus (including PMv and M1/S1), extending caudally into the supramarginal gyrus, and rostrally into the inferior frontal gyrus, pars opercularis and medially into the posterior insula.	L	–49	–8	24	239	
Cerebellum, lobules V and VI.	R	11	–64	–17	129	
Posterior superior temporal sulcus.	R	44	–38	11	54	
Caudal thalamus.	R	18	–33	6	46	

### 2.3. Discussion

In the first experiment we examined the pattern of functional activation in three PT regions (lateral, medial and caudal) during speech production with and without masking of auditory feedback to evaluate the contribution of endogenous (self-generated) and exogenous (masking noise) auditory input to PT activation. Here we used whispered speech in order to evaluate the effects of auditory feedback, which could be effectively masked through the use of masking noise. Under normal speech conditions, the use of masking noise only reduces the air-borne feedback signal with limited effect on bone-conducted feedback. The results of this study suggest, based on magnitude of activation in PT, that the left PT might be more strongly involved in processing speech than the right PT, although the activation patterns, overall, were similar. There was, however, evidence of functional heterogeneity within hemisphere. The ROIs differed in their response to the presence of auditory feedback. For the whispered speech conditions

(masked and not masked), the lateral PT was most strongly activated for auditory feedback when it was perceptible, that is when subjects were able to hear what they produced; in contrast, the medial PT exhibited a decrease in activation in the presence of perceptible feedback. Finally, the caudal PT was not modulated by the presence or absence of auditory feedback. When the masking noise was combined with the self-generated feedback there was no significant increase in activation magnitude in any PT region over and above the auditory input from an exogenous source (passive listening to masking noise). Hence, it appears that the PT, as a whole, does not merely sum auditory input but selectively responds to auditory input that is of functional relevance, in this case, self-generated feedback. In sum, our results demonstrate some degree of specialization within each PT region.

Interestingly, while our results suggest regional specializations within PT, they do not support the hypothesis of a functional specialization of the left PT in the processing of speech and language. We found similar activation patterns in



**Fig. 3 – Results of the ROI analyses for Experiment 1.** The vertical axis represents the difference between activation levels for the whispering conditions (Whisper–Whisper masked) expressed as percentage of BOLD signal change, for each of the left (A) and right (B) PT regions (medial, lateral, and caudal). The asterisks indicate significance (FDR corrected;  $q = .05$ ). The error bars represent the standard error of the mean (SE).

the left and right PT, with an overall greater activation magnitude in the left PT. In the following experiment we extend the investigation of potential regional and hemispheric specialization in PT during speech perception and production with a more naturalistic set of conditions comparing overt speaking (with voiced feedback) when elicited from written or auditory stimuli.

### 3. Experiment 2

#### 3.1. Materials and method

##### 3.1.1. Participants

Participants were ten healthy adults (mean  $27.3 \pm 3.8$  years; 5 females). All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971) and were native speakers of Canadian English. All participants had normal or corrected-to-normal vision, and reported no history of speech, language or learning difficulties. Participants were screened for any contraindication to MRI. Informed written consent was obtained from each participant. The study was approved by the MRRC and the MNIMNI Research Ethics Committee.

##### 3.1.2. Experimental procedures

The experiment consisted of six experimental conditions that were presented in pseudo-random blocks of 12 or 13 trials. The conditions included (1) silent word reading (RS), (2) reading aloud (RA), (3) auditory word listening (AWL), (4) AWR, (5) picture viewing and (6) picture naming. Each condition was repeated 50 times each. Only the first four tasks were analyzed for the present experiment, which resulted in a  $2 \times 2$  design: modality (visual, auditory) and task (passive, active).

The stimuli were 50 concrete words derived from a set of objects pictures of chosen from Snodgrass and Vanderwart (1980) (see Supplemental material for the list of all stimuli). The words were highly concrete (mean  $604 \pm 18$ ), highly imageable (mean  $599 \pm 25$ ) and highly familiar (mean  $560 \pm 37$ ). The mean written frequency of the words was  $40 \pm 29$  (range 10–127) (Francis and Kucera, 1982). The mean number of letter was  $4.6 \pm 1.2$  and the mean number of syllables was  $1.3 \pm .5$ .

Each block began with the presentation of a visual instruction, which was followed by the presentation of an auditory or a visual stimulus for 700 msec. A sparse sampling image acquisition protocol was used in which, at the end of each volume acquisition, a 2.36 sec period of silence occurred during which the gradients were switched off. The stimuli were presented, and the overt responses produced, during the silent interval. The subjects wore MR compatible headphones (Commander XG, Resonance technology, CA, USA) and their verbal responses were recorded through a MR compatible microphone attached to the headphones and digitized directly onto a laptop computer. The baseline condition, which was also presented in blocks of 12 or 13 trials, was a visual fixation condition during which participants stared at the word “rest”. All stimuli were presented through Presentation® software (Neurobehavioral Systems, Albany, CA, USA). The experiment lasted approximately 40 min.

##### 3.1.3. Image acquisition

The data were acquired on a 3 T Siemens Trio MR scanner (Siemens AG, Erlangen, Germany) at the MNI. Thirty-six axial slices (whole brain coverage) oriented parallel to the AC–PC line (thickness = 4 mm, no gap, FOV =  $256 \times 256$  mm<sup>2</sup>, matrix =  $64 \times 64$ ) were acquired in 2.64 sec using a multi-slice EPI sequence (TE = 50 msec, TR = 5 sec, delay in TR = 2.36 sec). The delay in TR occurred following each volume acquisition. The slices had a spatial resolution of  $4 \times 4 \times 4$  mm. Two experimental runs (14 min each) resulted in the acquisition of 382 T2\*-weighted BOLD images acquired in descending order. High-resolution T1-weighted volumes were acquired for anatomical localization (matrix  $256 \times 256$  mm, 160 slices,  $1 \times 1 \times 1$  mm, no gap, TE = 9.2 msec, TR = 22 msec). Participants' head was immobilized by means of a vacuum-bag filled with polystyrene balls and a forehead-restraining device (Hybex Innovations, St-Leonard, Qc, Canada).

##### 3.1.4. Image analysis

The same analysis flow was used in both studies. T-statistical images were computed for four contrasts: (1) AWL minus baseline; (2) AWR minus baseline; (3) RS minus baseline, and (4) RA minus baseline. A cluster correction for multiple comparisons across the brain was implemented in the AFNI program 3dClusterSim, which takes into account the amount of smoothing in the data, the number of voxels and their sizes. Based on 3dClusterSim results, we determined that a FWE rate of  $p < .05$  is achieved with a minimum cluster size of 44 contiguous voxels each significant at  $p < .01$ .

##### 3.1.5. ROI analysis

The process of identifying and extracting data in PT was identical to what was done in Experiment 1 with a manual

segmentation of PT into three non-overlapping medial, lateral and caudal regions. We then extracted the mean percentage of signal change for each condition and each PT region, as well as a number of difference scores. All analyses focused on (1) the difference between the two reading conditions (Read aloud–Read silently), (2) the two auditory conditions (AWR–AWL) and (3) the two speaking conditions (AWR–Read aloud). For each analysis, we conducted a 2-way ANOVA with repeated measurements, with factor ROI (Medial, Lateral, Caudal PT) and Hemisphere (left, right).

### 3.1.6. SNR analyses

The sensitivity of our experimental design was evaluated using the same procedure described in [Experiment 1](#).

## 3.2. Results

### 3.2.1. Behavioral data

No errors were made in the speaking tasks.

### 3.2.2. Neuroimaging results

**3.2.2.1. SNR ANALYSES.** The mean SNR for the group was 83.5 ( $\pm 10.3$  SD). Our simulations determined that, given a power of .80, the minimum SNR to detect a MR signal change of .5% was 59, whereas the minimum SNR to detect a MR signal change of 1% was 30. Thus, the design had sufficient SNR to detect significant comparisons of interest. These results are illustrated in [Supplementary Fig. S1](#).

**3.2.2.2. Voxel-based analyses.** As illustrated in [Fig. 4](#), for RA, AWL and AWR, the whole brain analyses revealed significant activation in the transverse temporal gyrus of Heschl and in the PT, bilaterally. For the two speaking conditions (RA and AWR), additional activity was found in the rostral and caudal regions of the precentral gyrus, bilaterally, in the IFG, also bilaterally, as well as in the right cerebellum (lobules V and VI) and in the left the rostral cingulate gyrus and sulcus. Activation in the insula was found during AWL, ARW and RA, but not RS. A list of all areas of significant activation, for each of the condition compared with baseline, is provided in [Table 2](#).

**3.2.2.3. ROI ANALYSIS.** In order to examine hemispheric and regional differences in PT activation patterns during speech production, we focused on three different contrasts: (1) Read Aloud compared to Read Silently (2) AWR compared to AWL and (3) AWR compared to Read Aloud. These contrasts isolate (1) the effect of overt versus covert word production, (2) the effect processing one's own auditory feedback over and above the processing of an external auditory stimulus, and, finally, (3) the difference between auditorily and visually triggered word production.

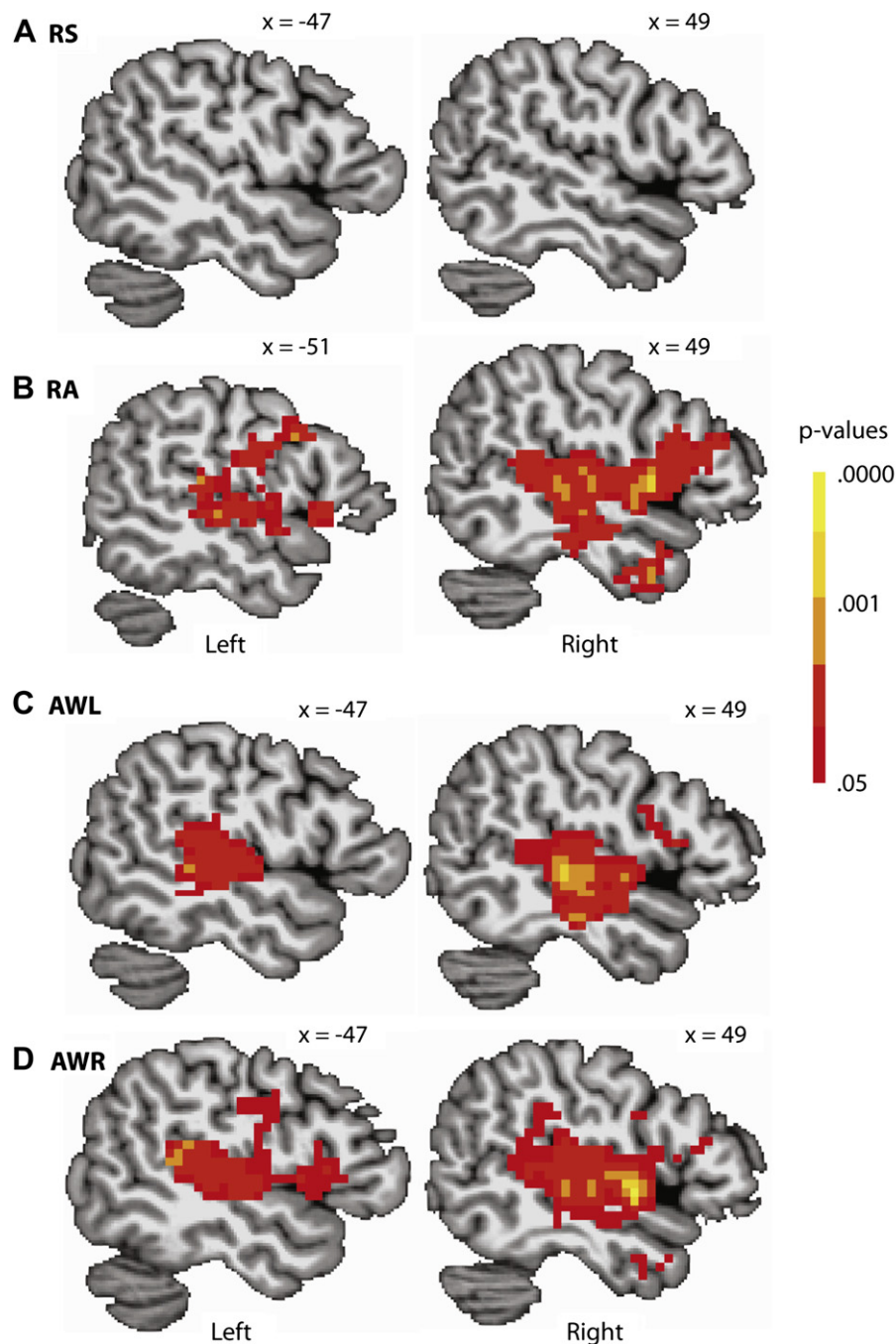
In a first step, we computed a set of difference scores removing the effect of Read silently from the effect of Read aloud (Read aloud–Read silently). These scores were entered in a 2-way ANOVA with repeated measurements with factor ROI (lateral PT, medial PT, caudal PT) and Hemisphere (left, right). Results show a main effect of ROI [ $F_{(2,9)} = 6.79, p = .006$ ], no main effect of Hemisphere [ $F_{(1,9)} = .03, p = .86$ ], and no interaction [ $F_{(2,18)} = .35, p = .71$ ]. As illustrated in [Fig. 5](#), two-

tailed FDR-corrected ( $i = 3; q = .05$ ) paired sample t-tests reveal that the caudal PT was significantly different from both the lateral PT [ $t_{(9df)} = 3.36, p = .016$ ], and from the medial PT [ $t_{(9df)} = 3.14, p = .033$ ]. While the lateral and medial PT exhibited a significant increase in activation for Read aloud compared to Read silently, the caudal PT showed no difference in activation. None of the regions was significantly active in the read silent condition; the caudal PT was not significantly active during Read silently or Read aloud.

In a second step, we computed a set of difference scores removing the effect of AWL from the effect of AWR. These scores were entered in a 2-way ANOVA with repeated measurements with factor ROI (lateral PT, medial PT, caudal PT) and Hemisphere (left, right). Results show no main effect of ROI [ $F_{(2,9)} = .02, p = .98$ ], no main effect of Hemisphere [ $F_{(1,9)} = .09, p = .77$ ], and no interaction [ $F_{(2,18)} = .39, p = .68$ ]. None of the difference scores was significantly different from zero, indicating no difference in activation magnitude between AWL and AWR. As shown in [Supplementary Fig. S3](#), inspection of the activation patterns reveals strong activation in both AWL and AWR.

Finally, in a third step, (1) we examined whether each of the ROIs was significantly active for both production tasks (Read aloud and AWR) ( $p \leq .05$ ), and (2) we computed a set of difference scores removing the effect of Read aloud from the effect of AWR. Results of the first analysis showed that the lateral and medial PT were significantly active, bilaterally, for both speaking conditions, while the bilateral caudal PT was only significantly active for AWR. For the second analysis, the difference scores were entered in a 2-way ANOVA with repeated measurements with factor ROI (lateral PT, medial PT, caudal PT) and Hemisphere (left, right). Results show a main effect of ROI [ $F_{(2,9)} = 7.05, p = .005$ ], no main effect of Hemisphere [ $F_{(1,9)} = .04, p = .84$ ], and no interaction [ $F_{(2,18)} = .19, p = .83$ ]. As illustrated in [Fig. 6](#), two-tailed FDR-corrected paired sample t-tests indicate that the lateral PT showed a significantly greater increase sensitivity for this contrast when compared to the medial PT [ $t_{(9df)} = 3.29, p = .016$ ] and the caudal PT [ $t_{(9df)} = 4.29, p = .033$ ].

**3.2.2.4. LINEAR TREND ANALYSES.** In order to further examine the contribution of the different sectors of PT to feedback processing, we conducted a trend analysis across study 1 and 2, which was warranted by the fact that the same ROIs were used across studies, the signal normalized to percentage of change, and the SNRs were comparable across studies. The focus of this analysis was to test for a progressive increase in activation in each PT region, for speaking tasks with increasing level of auditory feedback. For this analysis, hence, we examined a linear trend from Whisper masked (articulation, masked voiceless feedback), to Whisper (articulation and voiceless feedback), to Read aloud (articulation, voiced feedback). This analysis was conducted separately for the left and right PT regions.  $p$  values were FDR corrected. As shown in [Fig. 7](#), this analysis revealed that activity in the left and right lateral PT did not increase commensurate to an increase in the amount of auditory feedback [left:  $F_{(1,9)} = 4.37, p = .068$ ; right:  $F_{(1,9)} = 1.58, p = .24$ ], similar results were obtained for the caudal PT [left:  $F_{(1,9)} = .90, p = .77$ ; right:  $F_{(1,9)} = .098, p = .76$ ]. In contrast, in the medial PT, there was a significant linear trend on the left [ $F_{(1,9)} = 6.99, p = .027$ ] but not on the right [ $F_{(1,9)} = .60, p = .45$ ].



**Fig. 4 – Cluster-corrected group-level activation (FWE rate of  $p < .05$ , achieved with a minimum cluster size of 44 contiguous voxels each significant at  $p < .01$ ) from Experiment 2 for the contrast of each condition (RS, RA, AWL, AWR) against a resting baseline. Data are overlaid on axial views of a participant's T1-weighted MRI transformed into stereotaxic space using the MNI305 template. All coordinates are in MNI space.**

### 3.3. Discussion

The main focus of Experiment 2 was to extend and refine the results of Experiment 1, which suggested that different parts of PT participate differentially in speech production and possibly in speech perception. Overall, our results support and extend the results of Experiment 1 in showing regional differences in the processing of the auditory signal that accompany single word production in a variety of contexts.

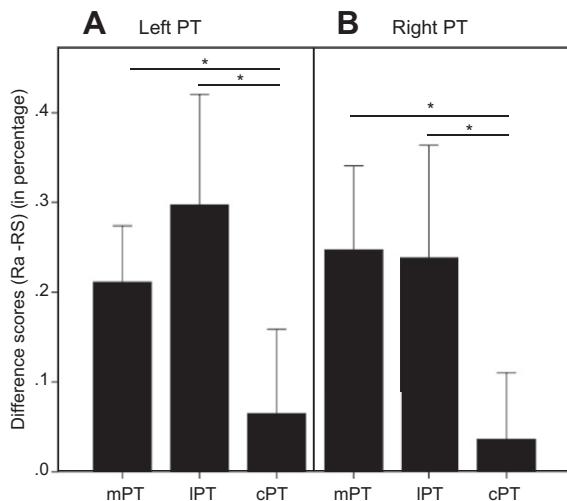
Importantly, our results demonstrate that PT, as a whole, is activated during passive listening, RA and AWR, but not during silent reading. Activation of PT during RA but not during reading silently indicates that PT is activated by endogenous (self-generated) auditory feedback consistent with the results from Experiment 1 using whispered speech. However, when comparing the activation associated with normal (voiced) speech and whispered speech it becomes clear that when activated by voiced feedback, the lateral and

**Table 2 – Table of statistically significant activation at the whole-brain level, corrected for multiple comparison (FWE rate of  $p < .05$ , achieved with a minimum cluster size of 44 contiguous voxels each significant at  $p < .01$ ), for Experiment 2, for each condition compared against a resting baseline: (1) Read silently, (2) Read aloud, (3) AWL and (4) AWR.**

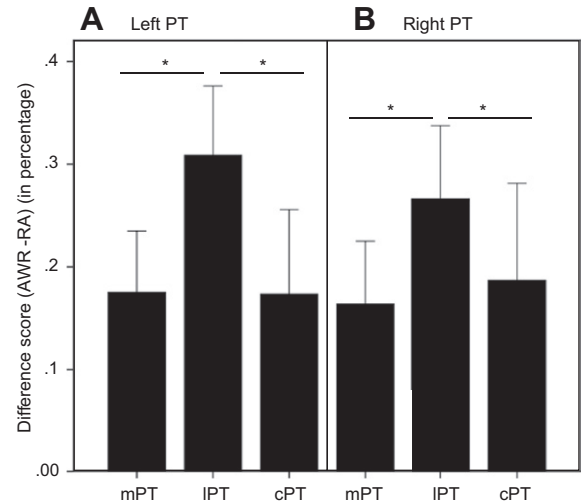
Description	Hemi	x	y	z	Voxels	p value
<b>A. Read silently</b>						
Lingual gyrus, striate cortex.	L	–8	–71	–2	44	.01000
<b>B. Read aloud</b>						
Transverse temporal gyrus and planum temporale, extending downward into the superior temporal sulcus, medially into the insula, and laterally and rostrally into the inferior frontal gyrus (pars opercularis and triangularis).	R	50	–13	5	300	.00027
Transverse temporal gyrus and planum temporale, extending into the superior temporal sulcus.	L	–43	–25	11	146	.00029
Caudal middle temporal gyrus.	L	–37	–49	–1	53	.00006
Para cingulate sulcus, and medial frontal gyrus (pre-SMA).	L	–8	14	39	49	.00002
<b>C. AWL</b>						
Transverse temporal gyrus and planum temporale, expanding medially into the posterior insula, and ventrally into the superior temporal sulcus and middle temporal gyrus.	R	53	–26	0	331	.00001
Transverse temporal gyrus and planum temporale, expanding medially into the posterior insula, and ventrally into the superior temporal sulcus.	L	–38	–26	–23	306	.00007
<b>D. AWR</b>						
Transverse temporal gyrus and planum temporale, superior temporal sulcus, expanding medially into the posterior insula and putamen and more dorsally and laterally on the pars opercularis of IFG, and on the precentral gyrus and central sulcus.	R	45	–17	4	673	.00000
Transverse temporal gyrus and planum temporale, superior temporal sulcus, expanding medially into the posterior insula, the putamen, thalamus.	L	–32	–23	5	422	.00001
Cerebellum lobules V and VI.	R	23	–53	–19	80	.00000
Posterior cingulate gyrus.	L	–9	–26	29	49	.00002

medial portions of PT are more responsive than the caudal portion. One finding of note related to word reading is the lack of PT activation (in all three regions) during silent reading (RS), a finding that cannot be attributed to a lack of sensitivity of

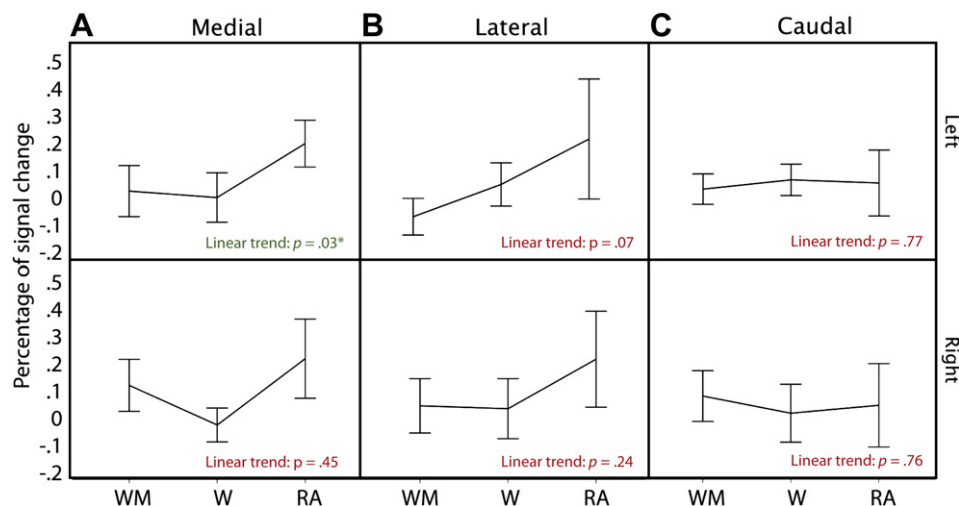
our experimental design, as revealed by the SNR analyses, or our spatial smoothing window (Buchsbaum et al., 2005). The lack of activation in the silent reading condition contrasts with previous studies employing silent reading yielding



**Fig. 5 – Results of the ROI analyses for Experiment 2. The vertical axis represents the difference between activation levels for the reading conditions (Read aloud–Read silent) expressed as percentage of BOLD signal change, for each of the left (A) and right (B) PT regions (medial, lateral, and caudal). Asterisks indicate significant differences between PT regions.**



**Fig. 6 – Results of the ROI analyses for Experiment 2. The vertical axis represents the difference between activation levels for the speaking conditions (AWR–Read aloud) expressed as percentage of BOLD signal change, for each of the left (A) and right (B) PT regions (medial, lateral, and caudal). Asterisks indicate significant differences between PT regions.**



**Fig. 7** – Linear trend analyses for the speaking tasks across Experiments 1 and 2. The vertical axis represents the mean percentage of BOLD signal change for the medial PT (A), lateral PT (B) and caudal PT (C), for the left PT (top row) and for the right PT (bottom row). Error bars represent the 90% confidence interval (CI) of the mean.

activation in the left posterior PT area identified as area Spt (Hickok et al., 2000, 2003; Okada et al., 2003; Buchsbaum et al., 2005; Pa and Hickok, 2008). In previous studies, however, silent reading was combined with verbal working memory or articulatory rehearsal requirements. In the present study, participants were simply asked to read silently each visually presented word, and similar to previous observations of silent reading (see Fiez and Petersen, 1998; Turkeltaub et al., 2002; Mechelli et al., 2003) no activation in PT was observed. This finding more clearly highlights the processing requirements needed to activate the left posterior PT during reading and suggests that sound–symbol associations and sensorimotor transformations may not be part of the silent reading process.

Another interesting result from Experiment 2 is the reduced level of activity in all regions of PT for word repetition compared to word listening. Given that re-reference provides a substantial level of activation in all regions of PT, no difference in activation magnitude suggests a modulation of auditory input during speech production, reflecting a speech-induced suppression of auditory cortex activity (Numminen et al., 1999; Curio et al., 2000; Houde et al., 2002; Poulet and Hedwig, 2002). The role of speech-induced suppression is unclear but it is most likely related to changing the sensitivity of auditory neurons to self-generated speech (Behroozmand et al., 2010; Eliades and Wang, 2008; Poulet and Hedwig, 2003). It appears that the different regions of PT exhibit both idiosyncratic as well as general processing features such as suppression.

## 4. General discussion

In the recent years, the role of the left PT in speech and language has received much attention (e.g., Hickok and Poeppel, 2000, 2004, 2007; Warren et al., 2005a, 2005b). Despite the renewed interest in this region, the details of the functional and anatomical organization of the human PT remain to be uncovered. The present study is the first, to our knowledge, to examine the contribution of three anatomically

distinct regions of the left and right PT to speech production using a set of four overt word production tasks with different levels of exogenous and endogenous auditory stimulation. Specifically, we examined the potential regional and inter-hemispheric specialization in PT for the production and perception of language. Our results demonstrate regional differences in the functional activation patterns in PT but similar inter-hemispheric trends. In the following paragraphs, we summarize and interpret these results.

### 4.1. Regional specializations in PT

Our results highlight important differences in the manner in which different regions of PT are modulated by different language tasks involving speech perception and speech production, and therefore emphasize the importance of considering the anatomical complexity in PT in interpreting functional magnetic imaging data (fMRI) data. Based on these findings, it appears that each region may be acting upon the incoming auditory signal in different ways and for different purposes. Here we suggest (1) that the lateral PT acts as a general auditory processing area, (2) that the medial PT is involved in processing speech feedback possibly as part of a feedback monitoring loop, and (3) that the caudal PT is involved in auditory–motor transformation for speech production.

#### 4.1.1. Lateral PT: general auditory processing

The lateral PT was the only region whose activation magnitude increased for whispered auditory feedback relative to masked feedback. The apparent differential response to the presence of masking noise, whether applied during passive listening or in combination with self-generated feedback, suggests a stronger or more prominent role for this part of PT in auditory processing. In addition, the comparison of Experiments 1 and 2 reveals (1) that activation in lateral PT increased linearly as the amount of auditory stimulation increased, and (2) that it was more sensitive than all other ROIs to the contrast of AWR and RA, which isolates the effect of adding external auditory

stimulation. Taken together, these findings suggest that the lateral PT is not specifically related to speech production or feedback control but instead in general auditory processing, perhaps acting, as suggested by Griffiths and Warren (2002), as a computational hub whose function is to disambiguate complex sounds, a process that would be achieved by isolating different properties of the acoustic objects and matching them to previously stored auditory templates.

#### 4.1.2. Medial PT processing speech feedback

The medial PT appears to be involved in processing self-generated speech possibly as part of a feedback monitoring loop. The left medial PT was the only region to show an increase in activation magnitude commensurate with an increase in the amount of feedback present. Compared to the lateral PT, it was less sensitive to the comparison of AWR and RA, suggesting a greater sensitivity to internal than external feedback. This interpretation is consistent with previous studies showing sensitivity of the medial PT to the presence of delayed (Takaso et al., 2010) and shifted auditory feedback (Tourville et al., 2008), which supports a role for medial PT in an online feedback control system.

#### 4.1.3. Caudal PT and auditory–motor transformation

For the caudal PT, the lack of a linear increase in activity with increasing amount of auditory feedback during speech production is consistent with previous results showing caudal PT/Spt activation for tasks not involving auditory feedback such as silent speech (Hickok et al., 2000, 2003; Buchsbaum et al., 2001; Wise et al., 2001; Huang et al., 2001; Papathanassiou et al., 2000; Okada et al., 2003; Shergill et al., 2002; Callan et al., 2006; Pa and Hickok, 2008). Moreover, the caudal PT was not activated for RA but only for AWR, a task requiring the conversion of an auditory target into a motor program, as well as for AWL. This suggests that this region is not involved in processing self-generated feedback per se, but it may be involved in monitoring external auditory input in order to facilitate internally-generated motor commands. This interpretation is in line with a recent fMRI study in which participants were asked to listen to songs and speech, as well as to speak and sing spontaneously and in synchrony with external auditory stimuli. Results showed that activation in PT was found only in the synchronized production tasks (Saito et al., 2006). In keeping with these and previous results, here we suggest that the caudal PT may be involved in the process of integrating auditory and motor signals for sensorimotor transformation for speech production (Hickok and Poeppel, 2000, 2004, 2007), or possibly articulatory rehearsal, which would account for the activation in this region during AWL.

#### 4.2. Inter-hemispheric patterns

As discussed in the Introduction, the potentially “special” role of the left PT to auditory language processing has been a central theme in cognitive neuroscience since the first discovery of a leftward asymmetry in this region in the nineteen sixties (Geschwind and Levitsky, 1968; Galaburda et al., 1978; Steinmetz and Galaburda, 1991). In the present study, we found a number of interesting regional differences in the processing of speech across the different PT regions.

Interestingly, our results demonstrate similar functional activation patterns, and functional activation magnitude, in the left and right PT regions, which suggests a bilateral processing of incoming auditory signal, which does not support the hypothesis of a distinct contribution of the left PT to the processing and production of speech. Admittedly, however, it is possible that some regions of the left PT contribute more strongly than their corresponding regions in the right PT during different tasks, for instance more demanding phonological tasks. This will have to be examined in future studies.

#### 4.3. Conclusion

In summary, the present findings stress the importance of considering the anatomical complexity in PT in interpreting fMRI data. Indeed, the functional activation patterns of three PT regions suggest important functional difference between the different regions of PT. The PT, as a whole, appears to be part of multiple networks in which internal and external auditory input is acted upon in multiple ways and for multiple purposes depending on the task.

#### Acknowledgments

We would like to thank Anthony S. Dick for his help with the analysis of SNR and his comments on several versions of this manuscript, the research staff and MR technicians at the Montreal Neurological Institute (Montreal, Canada) for their help in collecting the data, and all the participants. This study was supported by research grants from the Canadian Institute of Health Research (CIHR) and Natural Science and Engineering Council of Canada to V.L. Gracco and a CIHR post-doctoral fellowship to P. Tremblay.

#### Appendix. Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cortex.2011.09.004.

#### REFERENCES

- Behroozmand R, Liu H, and Larson CR. Time-dependent Neural Processing of Auditory Feedback during Voice Pitch Error Detection. *Journal of Cognitive Neuroscience*, 23(5): 1205–1217, 2010.
- Belin P, Zatorre RJ, and Ahad P. Human temporal-lobe response to vocal sounds. *Brain Research. Cognitive Brain Research*, 13(1): 17–26, 2002.
- Binder JR, Frost JA, Hammeke TA, Rao SM, and Cox RW. Function of the left planum temporale in auditory and linguistic processing. *Brain*, 119(Pt 4): 1239–1247, 1996.
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, and Prieto T. Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience*, 17(1): 353–362, 1997.
- Binder JR, Frost JA, Hammeke TA, Bellgowan PSF, Springer JA, Kaufman JN, and Possing ET. Human temporal lobe activation

- by speech and nonspeech sounds. *Cerebral Cortex*, 10(5): 512–528, 2000.
- Bohland JW and Guenther FH. An fMRI investigation of syllable sequence production. *NeuroImage*, 32(2): 821–841, 2006.
- Buchsbaum BR, Hickok G, and Humphries C. Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science*, 25: 663–678, 2001.
- Buchsbaum BR, Olsen RK, Koch PF, Koch PF, Kohn P, Kippenhan JS, and Berman KF. Reading, hearing, and the planum temporale. *NeuroImage*, 24: 444–454, 2005.
- Callan DE, Tsytsarev V, Hanakawa T, Callan AM, Katsuhara M, Fukuyama H, and Turner R. Song and speech: Brain regions involved with perception and covert production. *NeuroImage*, 31(3): 1327–1342, 2006.
- Caplan D, Gow D, and Makris N. Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. *Neurology*, 45(2): 293–298, 1995.
- Collins DL, Neelin P, Peters TM, and Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *Journal of Computer Assisted Tomography*, 18(2): 192–205, 1994.
- Cox RW. Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, 29: 162–173, 1996.
- Curio G, Neuloh G, Numminen J, Jousmaki V, and Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*, 9(4): 183–191, 2000.
- Dhanjal NS, Handunnetthi L, Patel MC, and Wise RJ. Perceptual systems controlling speech production. *The Journal of Neuroscience*, 28(40): 9969–9975, 2008.
- Dick AS, Goldin-Meadow S, Hasson U, Skipper JI, and Small SL. Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Human Brain Mapping*, 30(11): 3509–3526, 2009.
- Dorsaint-Pierre R, Penhune VB, Watkins KE, Neelin P, Lerch JP, Bouffard M, and Zatorre RJ. Asymmetries of the planum temporale and Heschl's gyrus: Relationship to language lateralization. *Brain*, 129(Pt 5): 1164–1176, 2006.
- Eden GF, Joseph JE, Brown HE, Brown CP, and Zeffiro TA. Utilizing hemodynamic delay and dispersion to detect fMRI signal change without auditory interference: The behavior interleaved gradients technique. *Magnetic Resonance in Medicine*, 41(1): 13–20, 1999.
- Edmister WB, Talavage TM, Ledden PJ, and Weisskoff RM. Improved auditory cortex imaging using clustered volume acquisitions. *Human Brain Mapping*, 7(2): 89–97, 1999.
- Eliades SJ and Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, 453: 1102–1106, 2008.
- Fiez JA and Petersen SE. Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences of the United States of America*, 95: 914–921, 1998.
- Francis WN and Kucera H. *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton Mifflin, 1982.
- Friederici AD, Meyer M, and von Cramon DY. Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language*, 75(3): 289–300, 2000.
- Fullerton BC and Pandya DN. Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. *The Journal of Comparative Neurology*, 504(5): 470–498, 2007.
- Galaburda AM, LeMay M, Kemper TL, and Geschwind N. Right–left asymmetries in the brain. *Science*, 4331: 852–856, 1978.
- Galaburda A and Sanides F. Cytoarchitectonic organization of the human auditory cortex. *The Journal of Comparative Neurology*, 190: 597–610, 1980.
- Geschwind N and Levitsky W. Human brain: Left–right asymmetries in temporal speech region. *Science*, 837: 186–187, 1968.
- Goulven J, Mazoyer B, Crivello F, and Tzourio-Mazoyer N. Left planum temporale: An anatomical marker of the left hemispheric specialization for language comprehension. *Cognitive Brain Research*, 18: 1–14, 2003.
- Gracco VL, Tremblay P, and Pike B. Imaging speech production using fMRI. *NeuroImage*, 26(1): 294–301, 2005.
- Griffiths TD, Buchel C, Frackowiak RS, and Patterson RD. Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, 1(5): 422–427, 1998.
- Griffiths TD and Warren JD. The planum temporale as a computational hub. *Trends in Neurosciences*, 25(7): 348–353, 2002.
- Hall DA and Plack CJ. Pitch processing sites in the human auditory brain. *Cerebral Cortex*, 19(3): 576–585, 2009.
- Hickok G and Poeppel D. Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4(4): 131–138, 2000.
- Hickok G, Erhard P, Kassubek J, Helms-Tillery AK, Naeve-Velguth S, Strupp JP, et al. A functional magnetic resonance imaging study of the role of left posterior superior temporal gyrus in speech production: Implications for the explanation of conduction aphasia. *Neuroscience Letters*, 287: 156–160, 2000.
- Hickok G, Buchsbaum B, Humphries C, and Muftuler T. Auditory–motor interaction revealed by fMRI: Speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience*, 15(5): 673–682, 2003.
- Hickok G and Poeppel D. Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1–2): 67–99, 2004.
- Hickok G and Poeppel D. The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5): 393–402, 2007.
- Hickok G, Okada K, and Serences JT. Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *Journal of Neurophysiology*, 101: 2725–2732, 2009.
- Hugdahl K, Bronnick K, Kyllingsbaek S, Law I, Gade A, and Paulson OB. Brain activation during dichotic presentations of consonant–vowel and musical instrument stimuli: A 15O-PET study. *Neuropsychologia*, 37(4): 431–440, 1999.
- Houde JF, Nagarajan SS, Sekihara K, and Merzenich MM. Modulation of the auditory cortex during speech: An meg study. *Journal of Cognitive Neuroscience*, 14: 1125–1138, 2002.
- Huang J, Carr TH, and Cao Y. Comparing cortical activations for silent and overt speech using event-related fMRI. *Human Brain Mapping*, 15: 39–53, 2001.
- Jacquemot C, Pallier C, LeBihan D, Dehaene S, and Dupoux E. Phonological grammar shapes the auditory cortex: A functional magnetic resonance imaging study. *The Journal of Neuroscience*, 23(29): 9541–9546, 2003.
- Jancke L, Specht K, Shah JN, and Hugdahl K. Focused attention in a simple dichotic listening task: An fMRI experiment. *Brain Research. Cognitive Brain Research*, 16(2): 257–266, 2003.
- Karbe H, Herholz K, Weber-Luxemburger G, Ghaemi M, and Heiss WD. Cerebral networks and functional brain asymmetry: Evidence from regional metabolic changes during word repetition. *Brain and Language*, 63(1): 108–121, 1998.
- Kumar S, Stephan KE, Warren JD, Friston KJ, and Griffiths TD. Hierarchical processing of auditory objects in humans. *PLoS Computational Biology*, 3(6): e100, 2007.
- Liberman AM and Mattingly IG. The motor theory of speech perception revised. *Cognition*, 21(1): 1–36, 1985.
- Lombard E. Le signe de l'élévation de la voie. *Annales des maladies de l'oreille, du larynx, du nez et du pharynx*, 37: 101–119, 1911, as cited in Lane H and Tranel B. *Journal of speech and hearing research*, 14: 677, 1971.

- Marshall JC. Planum of the apes: A case study. *Brain and Language*, 71(1): 145–148, 2000.
- Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, et al. The cortical representation of speech. *Journal of Cognitive Neuroscience*, 5(4): 467–479, 1993.
- Mechelli A, Gorno-Tempini M, and Price CJ. Neuroimaging studies of word and pseudoword reading: Consistencies, inconsistencies, and limitations. *Journal of Cognitive Neuroscience*, 15: 260–271, 2003.
- Numminen J, Salmelin R, and Hari R. Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters*, 265(2): 119–122, 1999.
- Oldfield RC. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1): 97–113, 1971.
- Okada K, Smith KR, Humphries C, and Hickok G. Word length modulates neural activity in auditory cortex during covert object naming. *NeuroReport*, 14(18): 2323–2326, 2003.
- Pa J and Hickok G. A parietal–temporal sensory–motor integration area for the human vocal tract: Evidence from an fMRI study of skilled musicians. *Neuropsychologia*, 46(1): 362–368, 2008.
- Papathanassiou D, Etard O, Mellet E, Zago L, Mazoyer B, and Tzourio-Mazoyer N. A common language network for comprehension and production: A contribution to the definition of language epicenters with PET. *NeuroImage*, 11(4): 347–357, 2000.
- Parrish TB, Gitelman DR, LaBar KS, and Mesulam MM. Impact of signal-to-noise on functional MRI. *Magnetic Resonance in Medicine*, 44(6): 925–932, 2000.
- Paus T, Perry DW, Zatorre RJ, Worsley KJ, and Evans AC. Modulation of cerebral blood flow in the human auditory cortex during speech: Role of motor-to-sensory discharges. *The European Journal of Neuroscience*, 8(11): 2236–2246, 1996.
- Peschke C, Ziegler W, Kappes J, and Baumgaertner A. Auditory–motor integration during fast repetition: The neuronal correlates of shadowing. *NeuroImage*, 47(1): 392–402, 2009.
- Petersen SE, Fox PT, Posner MI, Mintun M, and Raichle ME. Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, 331(6157): 585–589, 1988.
- Pfeifer RA. Pathologie der Hörstrahlung der corticalen Hörsphäre. In Bumke O and Forster O (Eds), *Handbuch der Neurologie*. Berlin: Springer, 1936: 533–626.
- Poulet JF and Hedwig B. A corollary discharge maintains auditory sensitivity during sound production. *Nature*, 418(6900): 872–876, 2002.
- Poulet JF and Hedwig B. Corollary discharge inhibition of ascending auditory neurons in the stridulating cricket. *The Journal of Neuroscience*, 23(11): 4717–4725, 2003.
- Praamstra P, Hagoort P, Maassen B, and Crul T. Word deafness and auditory cortical function. A case history and hypothesis. *Brain*, 114(Pt 3): 1197–1225, 1991.
- Riecker A, Brendel B, Ziegler W, Erb M, and Ackermann H. The influence of syllable onset complexity and syllable frequency on speech motor control. *Brain and Language*, 107(2): 102–113, 2008.
- Rivier F and Clarke S. Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: Evidence for multiple auditory areas. *NeuroImage*, 6(4): 288–304, 1997.
- Saito Y, Ishii K, Yagi K, Tatsumi IF, and Mizusawa H. Cerebral networks for spontaneous and synchronized singing and speaking. *NeuroReport*, 17(18): 1893–1897, 2006.
- Saur D, Kreher BW, Schnell S, Kümmerer D, Kellmeyer P, Magnus-Sebastian V, et al. Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, 105(46): 18035–18040, 2008.
- Schmahmann JD and Pandya DN. *Fiber Pathways of the Brain*. New York: Oxford University Press, 2006.
- Schmahmann JD, Pandya DN, Wang R, Dai G, D'Arceuil HE, de Crespigny AJ, and Wedeen VJ. Association fibre pathways of the brain: Parallel observations from diffusion spectrum imaging and autoradiography. *Brain*, 130(Pt 3): 630–653, 2007.
- Schnider A, Benson DF, Alexander DN, and Schnider-Klaus A. Non-verbal environmental sound recognition after unilateral hemispheric stroke. *Brain*, 117(Pt 2): 281–287, 1994.
- Shapleske J, Rossell SL, and Woodruff PW. The planum temporale: A systematic, quantitative review of its structural, functional and clinical significance. *Brain Research and Brain Research Review*, 29: 26–49, 1999.
- Shergill SS, Brammer MJ, Fukuda R, Bullmore E, Amaro Jr E, Murray RM, et al. Modulation of activity in temporal cortex during generation of inner speech. *Human Brain Mapping*, 16(4): 219–227, 2002.
- Schulz GM, Varga M, Jeffries K, Ludlow CL, and Braun AR. Functional neuroanatomy of human vocalization: An H215O PET study. *Cerebral Cortex*, 15(12): 1835–1847, 2005.
- Scheich H, Baumgart F, Gaschler-Markefski B, Tegeler C, Tempelmann C, Heinze HJ, et al. Functional magnetic resonance imaging of a human auditory cortex area involved in foreground–background decomposition. *The European Journal of Neuroscience*, 10(2): 803–809, 1998.
- Snodgrass JG and Vanderwart M. A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology. Human Learning and Memory*, 6(2): 174–215, 1980.
- Specht K and Reul J. Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: An auditory rapid event-related fMRI-task. *NeuroImage*, 20(4): 1944–1954, 2003.
- Spocter MA, Hopkins WD, Garrison AR, Bauernfeind AL, Stimpson CD, Hof PR, and Sherwood CC. Wernicke's area homologue in chimpanzees (pan troglodytes) and its relation to the appearance of modern human language. *Proceedings of the Royal Society B Biological Sciences*, 277: 2165–2174, 2010.
- Steinmetz H and Galaburda AM. Planum temporale asymmetry: In-vivo morphometry affords a new perspective for neuro-behavioral research. *Reading and Writing*, 3: 331–343, 1991.
- Sweet RA, Dorph-Petersen K-A, and Lewis DA. Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. *The Journal of Comparative Neurology*, 491: 270–289, 2005.
- Tanaka Y, Yamadori A, and Mori E. Pure word deafness following bilateral lesions. A psychophysical analysis. *Brain*, 110(Pt 2): 381–403, 1987.
- Takaso H, Eisner F, Wise RJ, and Scott SK. The effect of delayed auditory feedback on activity in the temporal lobe while speaking: A positron emission tomography study. *Journal of Speech, Language, and Hearing Research: JSLHR*, 53(2): 226–236, 2010.
- Tardif E and Clarke S. Intrinsic connectivity of human auditory areas: A tracing study with DiI. *The European Journal of Neuroscience*, 13(5): 1045–1050, 2001.
- Tourville JA, Reilly KJ, and Guenther FH. Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, 39(3): 1429–1443, 2008.
- Turkeltaub PE, Eden GF, Jones KM, and Zeffiro TA. Metaanalysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage*, 16: 765–780, 2002.
- von Economo C and Horn L. ber Windungsrelief, Masse und Rindenarchitektonik der Supratemporalfläche, ihre individuellen und ihre Seitenunterschiede. *Zeitschrift für Neurologie und Psychiatrie*, 130: 678–757, 1930.

- Wang P, Kjems U, Pedersen MS, Boldt JB, and Lunner T. Speech perception of noise with binary gains. *Acoustical Society of America*, 124(4): 2303–2307, 2008.
- Warren JD, Jennings AR, and Griffiths TD. Analysis of the spectral envelope of sounds by the human brain. *NeuroImage*, 24(4): 1052–1057, 2005.
- Warren JE, Wise RJ, and Warren JD. Sounds do-able: Auditory–motor transformations and the posterior temporal plane. *Trends in Neurosciences*, 28(12): 636–643, 2005.
- Wilson SM, Molnar-Szakacs I, and Iacoboni M. Beyond superior temporal cortex: Intersubject correlations in narrative speech comprehension. *Cerebral Cortex*, 18: 230–242, 2008.
- Wise R, Chollet F, Hadar U, Friston K, Hoffner E, and Frackowiak R. Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain*, 114(Pt 4): 1803–1817, 1991.
- Wise RJ, Scott SK, Blank SC, Mummery CJ, Murphy K, and Warburton EA. Separate neural subsystems within ‘Wernicke’s area’. *Brain*, 124(Pt 1): 83–95, 2001.
- Zatorre RJ, Evans AC, Meyer E, and Gjedde A. Lateralization of phonetic and pitch discrimination in speech processing. *Science*, 256(5058): 846–849, 1992.
- Zatorre RJ, Evans AC, and Meyer E. Neural mechanisms underlying melodic perception and memory for pitch. *The Journal of Neuroscience*, 14(4): 1908–1919, 1994.
- Zheng ZZ, Munhall KG, and Johnsrude IS. Functional overlap between regions involved in speech perception and in monitoring one’s own voice during speech production. *Journal of Cognitive Neuroscience*, 22(8): 1770–1781, 2010.