

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

Issue: *The Year in Cognitive Neuroscience***The anatomy of language: a review of 100 fMRI studies published in 2009**

Cathy J. Price

Wellcome Trust Centre for Neuroimaging, UCL, London, UK

Address for correspondence: Cathy Price, Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, 12, Queen Square, London WC1N 3BG, UK. c.price@fil.ion.ucl.ac.uk

In this review of 100 fMRI studies of speech comprehension and production, published in 2009, activation is reported for: prelexical speech perception in bilateral superior temporal gyri; meaningful speech in middle and inferior temporal cortex; semantic retrieval in the left angular gyrus and pars orbitalis; and sentence comprehension in bilateral superior temporal sulci. For incomprehensible sentences, activation increases in four inferior frontal regions, posterior planum temporale, and ventral supramarginal gyrus. These effects are associated with the use of prior knowledge of semantic associations, word sequences, and articulation that predict the content of the sentence. Speech production activates the same set of regions as speech comprehension but in addition, activation is reported for: word retrieval in left middle frontal cortex; articulatory planning in the left anterior insula; the initiation and execution of speech in left putamen, pre-SMA, SMA, and motor cortex; and for suppressing unintended responses in the anterior cingulate and bilateral head of caudate nuclei. Anatomical and functional connectivity studies are now required to identify the processing pathways that integrate these areas to support language.

Keywords: speech; comprehension; production; fMRI; language

Introduction

The aim of this paper was to provide an up to date and integrated review of functional imaging studies of language published in 2009. I included studies that (i) used fMRI, (ii) were published online between January and mid-October 2009, and (iii) were concerned with the functional anatomy of speech comprehension and production in the healthy adult brain. This excluded papers that used MEG, EEG, TMS, or optical imaging; that were published in 2008 or before; or that focused on populations with acquired or developmental language disorders. I also excluded papers on orthographic processing of written words, sign language, and multilingualism: each of these topics requires its own review.

Having selected more than 100 fMRI papers on speech comprehension or production, I categorized each paper according to the conclusions provided in their abstracts. The first categorization was *speech comprehension* or *speech production*. Within each of these main categories, I progressively grouped

the papers according to whether their stimuli were *prelexical*, *single words*, or *sentences*. This process illustrated the wide range of language topics that are currently being investigated with fMRI. Although all papers published prior to 2009 were excluded, my assumption was that the 2009 papers would have incorporated the most salient conclusions from former years, as well as assimilating their literature. Each of the 2009 papers listed in the references of this review article should therefore provide a list of related papers published prior to 2009.

Given the number of language fMRI studies published in 2009, it was not possible to do justice to the contribution of each one. Furthermore, a review is limited to the reviewer's own understanding of the topic and how the conclusions of each paper fit together. What follows is the overall picture that I saw when reviewing the selected literature. It emphasizes consistent rather than atypical results. Only time and more studies will determine whether the conclusions drawn are correct.

Anatomical terms

Across studies, different anatomical terms have been used for the same region. For example, the same inferior frontal region has been variously referred to as Broca's area, BA44, or pars opercularis. The anatomical distinction between anterior and posterior superior temporal lobe activation also varies across studies; and the same left temporo-parietal area has been referred to as the supramarginal gyrus (SMG), planum temporale, Spt (Sylvian parieto-temporal), superior temporal gyrus (STG), and the inferior parietal lobe (IPL). I have therefore tried to use a more consistent terminology (see Fig. 1), reported the MNI or Talairach coordinates $[x,y,z]$ of the effects I am referring to and estimated the subregion being referred to using both MNI and Talairach templates. I divided inferior frontal activations into four different regions: pars triangularis (pTr), pars orbitalis (pOr), dorsal pars opercularis (dpOp), and ventral pars opercularis (vpOp). In the superior temporal gyrus, anterior and posterior parts were in relation to Heschl's gyrus (HG).

Given the limited spatial resolution of fMRI and the use of different spatial normalization templates, it was not always possible to determine the location of reported coordinates, particularly when they fell at the border between two regions. For example, the lower bank of the sylvian fissure, posterior to HG is in the posterior planum temporale (pPT), and the upper bank of the same part of the sylvian fissure is in the ventral supramarginal gyrus (vSMG). These regions will not have the same connectivity or functions but it was difficult to tell when an activation fell in the lower or upper bank of the fissure because of their close proximity. One option is to refer to activation around the posterior part of the Sylvian fissure as Spt (Sylvian parieto-temporal), as in Hickok and colleagues.¹ However, this de-emphasizes the fact that the lower and upper banks of the sylvian fissure will have different functions. I have therefore labeled activation around Spt as pPT/vSMg to emphasize the convergence of two different regions without implying that they have the same function.

To provide summaries of the results, I marked the approximate location of the reported coordinates on a two-dimensional template of the brain. This further reduced the spatial precision of the results. Remarkably, however, the low-resolution cartoon illustrations were sufficient to highlight dis-

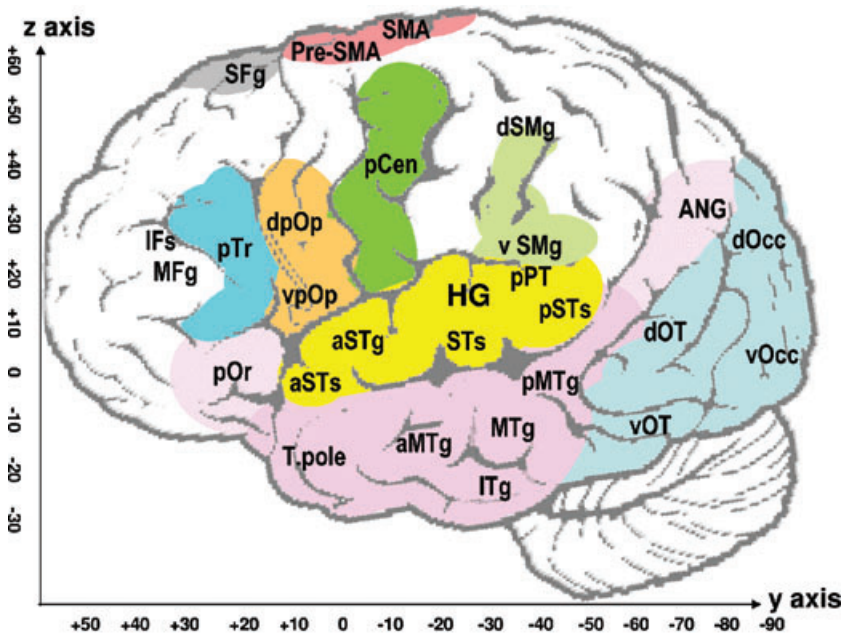
tinct spatial dissociations in the anatomical systems that support speech comprehension and production at the prelexical, word, and sentence levels (see Fig. 2). Future studies that dissociate activations within subject, rather than across study, are now needed to provide more accurate anatomical localization; and functional connectivity analyses with high temporal resolution are needed to determine how the widely distributed sets of regions interact with one another.

Review of speech comprehension

Auditory speech processing extracts meaningful information from continuously changing acoustic inputs. Many acoustic features predict the perceptual identity (intelligibility) of a given auditory stimulus including the temporal dynamics of the timbre, pitch, and volume in the auditory waveform and the integration of this information over short- and long-time frames. The type of information that is extracted will depend on the expectations and intentions of the listener. For example, auditory input can be translated into meaningful messages (semantics), articulatory associations (speech production), spellings, and physical actions (locomotion) toward or away from the source of the sound (*i.e.*, spatial processing).

As auditory speech information is ambiguous, comprehension can be facilitated by auditory cues from prosody and visual cues from concurrent hand gestures, lip movements, facial expressions, and writing. These additional cues are particularly useful under adverse listening conditions. For example, co-speech hand gestures provide semantic cues relevant to the intended message. Lip reading (viewing mouth movements) provides articulatory information and writing facilitates both semantic and articulatory processing. Therefore, speech comprehension, even at the single-word level, involves the integration of multiple cues from auditory and visual sources.

Speech comprehension at the sentence level is also ambiguous because the meaning of a sentence can be different from its parts, particularly for a metaphorical sentence such as "*You are the sun in my sky.*" Sentence comprehension is therefore constrained by our prior knowledge of the world (semantic constraints), what we are expecting to hear (context effects), and how the words are typically combined



- vOCC/dOCC:** ventral and dorsal occipital (approx. BA18/19)
vOT/dOT : ventral and dorsal occipito-temporal (approx. BA37)
aMTg/pMTg: anterior and posterior middle temporal gyrus (approx. BA21)
ITg: inferior temporal gyrus (approx. BA 20)
T. pole : temporal pole (approximately BA38)
HG: heschl's gyrus
STg: superior temporal gyrus (approx. BA 22)
aSTs/pSTs anterior and posterior superior temporal sulcus
pPT : planum temporale, posterior to HG (approx. BA42)
vSMg/dSMg: ventral and dorsal supramarginal gyrus (approx. BA40)
ANG : angular gyrus (approx. BA39)
pCen : precentral gyrus (BA 6 anteriorly, BA 4 posteriorly)
SMA: supplementary motor cortex
SFg: superior frontal gyrus
IFs: inferior frontal sulcus
MFg: middle frontal gyrus (approx. BA 46)
pOr: pars orbitalis (approx. BA 47)
pTr : pars triangularis (approx. BA 45);
vpOp/dpOp: ventral and dorsal pars opercularis (approx. BA44)

Figure 1. Anatomical terms. The names and location of the regions referred to in the text are illustrated on a cartoon sketch of the left hemisphere. This low-resolution sketch was deliberately chosen for all figures because it had minimal anatomical details, which were not appropriate in the context of the poor spatial precision of the activation loci illustrated in Figure 2, see introductory text for further explanation.

(*i.e.*, syntactic constraints/grammatical rules). As the meaning of the sentence emerges over time, auditory short-term memory is also required to hold individual words in memory until a satisfactory

interpretation of the word combination has been achieved. Each stage in the speech-processing hierarchy therefore involves the integration of bottom-up processing from auditory and visual inputs with

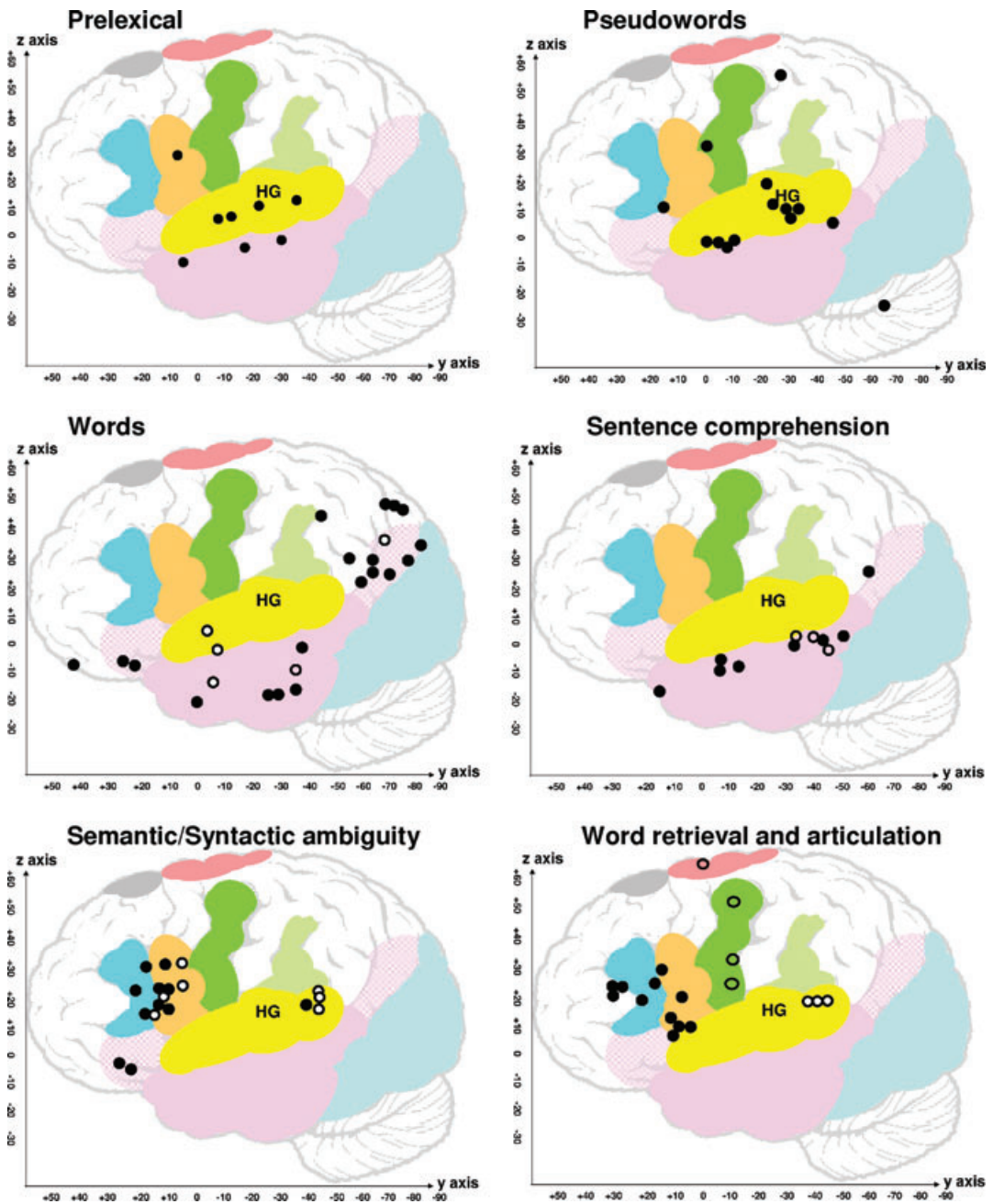


Figure 2. Summary of left hemisphere activation foci. The approximate location of the left hemisphere activation foci reported in the text. Right hemisphere foci were less consistent and not summarized here. Despite the poor spatial precision, the relative distributions of the different types of effects are valid and informative. *Top left panel:* prelexical processing^{3–5,7,8}; *top right panel:* pseudowords¹³; *middle left panel:* black circles = words relative to pseudowords,¹³ white circles = words relative to rotated words,^{9,10} open circles = effect of familiarity with auditory objects¹⁴; *middle right panel:* black circles = comprehension versus incomprehensible sentences,^{19,21,23,24} open circles = sentences with versus without visual gestures.^{27–29} *Lower left panel:* black circles = semantic ambiguity,^{9,19,35–39} white circles = syntactic ambiguity^{10,32,49,50,55}; *lower right panel:* black circles = word retrieval,^{43,76,79,81–86} open circles = overt articulation,^{83,87,96} white circles = covert articulation.^{72,73,84}

top-down predictions from our prior knowledge of language and the context in which it is being used.

My review of the anatomy of speech comprehension is divided into seven sections. These focus on prelexical auditory processing (prelexical phonemic processing); semantic recognition at the single-word level (semantic processing of spoken words); semantic recognition at the sentence level (sentence comprehension); semantic constraints at the sentence level (semantic constraints in sentence comprehension); syntactic constraints at the sentence level (syntactic constraints); the role of subvocal articulation during speech comprehension (subvocal articulation during speech comprehension), and the influence of prosodic information (the role of prosody in speech comprehension). I then (“Summary of activation related to speech comprehension”) sum up, and (“Left lateralization for speech comprehension”) hypothesize that top-down processing has a strong influence on determining left lateralization for language.

Prelexical phonemic processing

Prelexical processing occurs prior to the semantic recognition of the auditory stimulus. It includes acoustic analysis of the frequency spectrum and the integration of auditory features over different time frames. The result is phonemic categorization of the speech signal and the morphological parsing of word stems and inflectional affixes.

Rauschecker and Scott² review evidence that auditory inputs are processed along an antero-lateral gradient that starts at HG, progressing in the anterior direction toward the temporal pole. As auditory information progresses anteriorly and laterally, lower level auditory features are combined into increasingly complex and intelligible sounds. For example, middle and anterior parts of the superior temporal sulci respond to prelexical stimuli, such as consonant–vowel syllables, and there may even be subregions selective for particular speech-sound classes, such as vowels.

Further evidence for vowel selective responses has been presented in two recent studies.^{3,4} Leff and colleagues³ compared activation for hearing familiar vowels versus single formants and tones. The formants had the same intonation pattern as the vowels but they were less acoustically complex because they lacked stop bursts or other formants. Relative to the formants and tones, vowel stimuli increased activa-

tion in lateral left HG, anteriorly [−62 −12 6] and posteriorly [−64 −24 10]. Concurrent activation was also observed in the right superior temporal sulcus [−64 −10 −8]. Likewise, Britton and colleagues⁴ identified activation in lateral and anterior HG [−64 −7 5/59 −7 5] for hearing vowels versus tones. These bilateral superior temporal activations are likely to reflect differences in the acoustic complexity of vowel stimuli versus tone and formant stimuli.

To probe more speech selective responses, Leff and colleagues³ manipulated the acoustic quality of their stimuli within category (*i.e.*, vowel, formant, or tone). Using a mismatch paradigm, the same stimulus was repeated for a few trials followed by a new stimulus in the same category. In this context, the new stimulus results in an activation change that can be related to the processes that differ between the new stimulus and the prior stimuli. The results revealed a left-lateralized activation in the anterior superior temporal lobe [−56 6 −10] when the vowel changed and this effect was not observed for formant or tone changes. Myers and colleagues⁵ also used a mismatch paradigm, repeating a phonemic stimulus and measuring activation in response to acoustic deviants. They found left dorsal pars opercularis activation at [−41 8 27] when the phonetic category of the new stimulus was different versus the same as its predecessors. These mismatch responses^{3,5} can be explained in terms of “prediction error” when top-down expectations from prior experience fail to predict the bottom-up inputs.⁶ According to this account, left anterior temporal activation for vowel deviants³ and left dorsal pars opercularis activation for phonetic deviants⁵ reflect the influence of top-down predictions, with the location of the mismatch response depending on the type of stimulus and the type of prediction.

Specht and colleagues⁷ also investigated speech selective responses by manipulating acoustic complexity within category. Using a sound morphing technique, white noise was gradually changed into either prelexical speech (consonant-vowels) or music (piano or guitar chords). Bilateral activation in the anterior–superior temporal sulci at [−51 −3 −9/60 −6 −3] was observed irrespective of whether speech or music emerged but left lateralized speech selective responses were identified more posteriorly [−56 −36 12] and ventrally [−54 −18 −6]. Likewise, Vaden and colleagues⁸ report an effect

of phonological priming in superior temporal regions [$-63 -30 \ 3/45 -33 -3$] that were posterior and ventral to HG. Together these results illustrate speech selective responses in the posterior and ventral processing directions.

To summarize, prelexical processing of auditory speech increased activation in bilateral superior temporal gyri but left lateralized responses were observed when the noise merged into speech versus music⁷ or when there was a mismatch between the stimulus presented and the stimulus expected.^{3,4} In both these contexts, left lateralization may be driven by top-down predictions from prior experience rather than bottom-up processing. For example, the emergence of speech over time allows the participant to predict the morphing of the stimulus; and mismatch responses have been associated with the failure of top-down processing to predict the stimulus.⁶

The location of all left hemisphere coordinates reported above for prelexical speech versus non-speech processing are illustrated in the top left panel of Figure 2. This illustrates that prelexical activation is most extensive in the direction anterior to HG. However, the results do not support a straightforward gradient of increased speech selectivity as information progresses more anteriorly. Speech selectivity was observed in the ventral^{7,8} and posterior⁷ directions as well as the anterior direction.³ This suggests that there are multiple prelexical processing pathways, with the level of activation in each depending on the stimulus, task and top-down expectations from prior experience.

Semantic processing of spoken words

Semantic processing of auditory speech links heard sounds to what we know about their meanings. This has been probed by comparing familiar words to (i) closely matched but unintelligible *spectrally rotated words*^{9,10} and (ii) pseudowords that are intelligible but without meaning.^{11–13} A study of auditory familiarity, when the acoustic stimulus is held constant,¹⁴ is also reported in this section.

When participants listen to spoken sentences, relative to spectrally rotated speech (which is unintelligible), bilateral mid-to-anterior superior temporal activation has been reported by Friederici and colleagues¹⁰ at [$-58 -4 \ 4/62 -4 -14$] and by Oleser and Kotz⁹ at [$-60 -8 -6/62 -6 -4$] with the latter

authors⁹ also reporting activation in the left angular gyrus [$-48 -64 \ 38$]. It is important to note, however, that the bilateral mid-to-anterior superior temporal activations observed for familiar words relative to spectrally rotated speech are in very close proximity to those reported as common to music and prelexical speech by Specht and colleagues⁷ at [$-51 -3 -9/60 -6 -3$]. It may therefore be the case that the activation reported for the comparison of speech to spectrally rotated speech^{9,10} reflects prelexical differences between speech and rotated speech rather than differences in semantic processing. In contrast, semantic retrieval involves more distributed regions. For example, Sharp and colleagues¹⁵ have shown that when attention is directed to the meaning of single words versus the acoustic properties of spectrally rotated speech, activation is observed in a left-lateralized network including regions in the inferior temporal gyrus, anterior fusiform, hippocampus, angular gyrus, pars orbitalis, superior and middle frontal gyri, and the right cerebellum.

The comparison of spoken words to pseudowords provides a tighter control for prelexical processing and speech intelligibility than the comparison of spoken words to spectrally rotated speech. I found three studies^{11–13} published on line in 2009 that compared spoken words and pseudowords in the context of deciding if the stimulus was a word or not (*i.e.*, a lexical decision task). In a review of prior studies, Davis and Gaskell¹³ report an impressive double dissociation in activation for familiar words versus pseudowords with activation for words distributed in bilateral anterior middle temporal cortices, posterior temporal parietal cortices, and the precuneus, with left-lateralized activation in the temporal pole, posterior middle temporal cortex, anterior fusiform, pars orbitalis, middle frontal cortex, anterior cingulate and putamen, and the right precentral gyrus. The left hemisphere foci of these results are illustrated in the left middle panel of Fig. 2). This highlights word activations that are outside the areas associated with prelexical processing, extending further in the anterior, ventral and posterior directions, consistent with the results of the prelexical studies discussed in section “Pre-lexical phonemic processing” that suggest multiple speech-processing routes.

The meta-analysis of auditory words versus pseudowords reported by Davis and Gaskell¹³ also

demonstrated increased activation for pseudowords relative to words within the superior temporal areas associated with prelexical processing in the previous section (compare top left and top right panels in Fig. 2). Together the results suggest that pseudowords increased the demands on prelexical processing whereas real words increased activation in semantic areas. Indeed, many of the areas associated with word versus pseudoword processing¹³ have also been associated with semantic processing of written words and pictures of objects, as summarized in a review of 120 functional imaging studies by Binder and colleagues.¹⁶ Diaz and McCarthy¹⁷ have also reported greater anterior middle temporal lobe activation bilaterally at $[-59 -14 -27/66 -10 -21]$ for written content words (with high semantic associations) versus written function words (that carry minimal or ambiguous semantic associations).

The second study that compared spoken words and pseudowords during lexical decision was reported by Kotz and colleagues.¹² Each target stimulus (word or pseudoword) was primed with a rhyming or nonrhyming word or pseudoword. Rhyming pseudowords compared to rhyming words increased activation in bilateral superior temporal gyri as reported by Davis and Gaskell,¹³ whereas rhyming words compared to rhyming pseudowords increased activation in frontal and parietal regions. Kotz and colleagues¹² focus their interpretation on the observation that left inferior frontal activation at the level of the left pars orbitalis/pars triangularis $[-56 28 6]$ was greater for words than pseudowords.

The third study to compare lexical decisions to words and pseudowords was reported by Kouider and colleagues¹¹ who report higher activation for words than pseudowords in the precuneus, medial superior frontal gyrus, and the anterior cingulate. Although Kouider and colleagues¹¹ do not present any coordinates for these effects, it is interesting to note that the same regions were also reported earlier for words versus pseudowords¹³ and for phonological repetition of words.⁸ Activation in the precuneus and superior frontal gyrus have also been associated with semantic decisions on heard words,^{15,16} narratives,¹⁸ metaphors,¹⁹ and semantic versus phonological word generation.²⁰ In brief, activation for spoken words relative to pseudowords during lexical decision was identified in frontal and parietal areas that are also activated during semantic decision tasks.

Finally, an alternative approach for identifying semantic recognition of auditory stimuli is to keep the stimulus (and prelexical processing) constant and compare responses to the same stimuli before and after learning. Leech and colleagues¹⁴ used a video game to train participants to associate novel acoustically complex, artificial nonlinguistic sounds to visually presented aliens. After training, left posterior superior temporal activation at $[-54 -37 -1]$ increased with how well the auditory categories representing each alien had been learnt. The authors emphasize that part of what makes speech special is the extended experience that we have with it. This could include acoustic familiarity, enhanced audio-visual associations, or auditory memory.

To summarize this section, semantic processing of familiar auditory stimuli activates a distributed set of regions that surround the ventral, anterior, and posterior borders of the perisylvian regions supporting prelexical auditory speech processing (see middle left panel vs. top left panel in Fig. 2). The extended distribution of semantic activations suggests that there are numerous pathways supporting speech perception and comprehension (see Rauschecker and Scott²). This might be envisioned as activity spreading out and around a spider's web-like network, centered on HG.

Sentence comprehension

To tap speech comprehension at the sentence level, activation has been compared for hearing grammatically correct sentences with plausible versus implausible meanings. This comparison is expected to reveal semantic representations at the level of word combinations while controlling for phonological and lexical familiarity, syntactic processing, and working memory. Activation in four key regions has been reported: anterior and posterior parts of the left middle temporal gyrus,^{9,19,21} bilateral anterior temporal poles,^{21,22} left angular gyrus,^{9,19} and the posterior cingulate/precuneus.^{18,19} Each of these areas was also associated with semantic processing in the review of 120 studies reported by Binder and colleagues.¹⁶ Surprisingly, however, activation for plausible relative to implausible sentences is not as extensive as that associated with semantic processing of single words (see left and right middle panels in Fig. 2) but was primarily focused in anterior and posterior parts of the middle temporal gyrus, just ventral to the superior temporal sulci.

In the left anterior middle temporal locus, activation has been reported by Mashal and colleagues¹⁹ at $[-46 -6 -10]$ for semantically plausible more than implausible sentences; by Adank and Devlin²³ at $[-58 -8 -4]$ when time compressed auditory sentences were understood versus not understood during a sentence verification task; by Devauchelle and colleagues²⁴ at $[-60 -12 -8/60 -8 -12]$ when auditory sentence processing was primed by a sentence with the same meaning; by Obleser and Kotz⁹ at $[-52 -6 -14]$ for sentences with meanings that were difficult versus easy to predict; and by Hubbard and colleagues²⁵ at $[-57 -12 -8]$ when spoken speech was accompanied by beat gestures (rhythmic beating of hand) that enhance semantic meaning by providing intonation. These sentence-level comprehension effects are not specific to auditory words because anterior middle temporal activation has also been reported by Snijders and colleagues²⁶ at $[-56 -6 -16]$ for written sentences compared to unrelated word sequences; and by Kircher and colleagues²⁷ at $[-52 4 -8]$ when spoken sentences with abstract meanings were accompanied by visual gestures relative to speech only or gesture only. This suggests that the left anterior superior temporal sulcus is involved in multimodal sentence processing.

Likewise, left posterior middle temporal activation has been reported at $[-54 -41 -1]$ by Mashal and colleagues,¹⁹ for semantically plausible versus implausible sentences; more anteriorly at $[-41 -31 0]$ by Rogalsky and Hickok²¹ when meaningful sentences were compared to lists of unrelated words; and more posteriorly at $[-55 -52 2]$ by Adank and Devlin²³ when time compressed auditory sentences were understood versus not understood during a sentence verification task. Activation in the left posterior middle temporal cortex is also increased when auditory sentences are accompanied by visual observation of the speaker's body movements. For example, Holle and colleagues²⁸ reported activation at $[-48 -39 3]$ when participants viewed a speaker describing manual object actions (e.g., "now I grate the cheese") while making meaningful manual gestures versus hearing speech only or seeing gestures alone. This effect was stronger in adverse listening conditions that benefit most from visual cues. Likewise, Kircher and colleagues²⁷ found activation at $[-63 -43 -1]$ when spoken sentences with abstract meanings were accompanied by visual cues from gestures versus hearing speech only or seeing ges-

tures only; and Robins and colleagues²⁹ found activation at $[-66 -31 4]$ when participants viewed the face and emotional expression of a person speaking sentences versus hearing the sentence or viewing the face alone. Seeing the lip movements of the speaker also provides phonological cues,³⁰ therefore it is not surprising that the activation reported by Robins and colleagues²⁹ for speech and faces extended dorsally $[-45 -37 10]$ into the superior temporal gyrus.

The findings of Dick and colleagues³¹ further suggest that left posterior temporal activation in response to speech (spoken stories) and gesture reflects two sources of semantic information rather than an active integration process. Specifically, they found that perceiving hand movements during speech increased activation in the left posterior temporal regions irrespective of the semantic relationship between the hand movement and the accompanying speech. Thus, it did not depend on whether speech and gesture could be integrated into the same meaning or not. Instead, the authors conclude that listeners attempt to find meaning, not only in the words that speakers produce, but also in the hand movements that accompany speech.

The role of the left and right temporal poles in sentence processing relative to unrelated lists of words has been reported by Rogalsky and Hickok²¹ at $[-47 17 -18]$ and $[52 18 -20]$. Snijders and colleagues²⁶ also found bilateral temporal pole regions at $[-53 18 -30/54 20 -32]$ for written sentences compared to unrelated word sequences but the coordinates are more ventral than those reported by Rogalsky and Hickok.²¹ To distinguish auditory and visual sentence processing requires a within subjects comparison of (i) auditory sentences with plausible versus implausible meanings with (ii) visual sentences with plausible versus implausible meanings. All four of these conditions were included in a study reported by Richardson and colleagues,³² which found no interaction between plausible versus implausible sentences presented in the auditory versus visual modality. Likewise, a meta-analysis of semantic and sentence processing activation by Visser and colleagues²² found anterior temporal lobe activation for both auditory and visual sentences and these authors discuss the role for the anterior temporal cortex in amodal combinatorial semantics.

Left angular gyrus activation has been less consistently reported during sentence processing but was observed by Mashal and colleagues¹⁹ at $[-47 -59 25]$ for sentences with plausible versus implausible meanings; and by Obleser and Kotz⁹ at $[-46 -64 38]$ for the comparison of heard sentences to unintelligible spectrally rotated speech. Obleser and Kotz⁹ suggest that activation in the angular gyrus facilitates sentence comprehension via top-down activation of semantic concepts. Likewise, Carreiras and colleagues³³ demonstrated a top-down role for the angular gyrus at $[-48 -74 28]$ during reading relative to object naming, and Brownsett and Wise³⁴ highlight the role of the left angular gyrus in both speaking and writing. In the medial parietal lobes, Whitney and colleagues¹⁸ highlight the importance of the right precuneus and bilateral posterior/middle cingulate cortices for narrative language comprehension, associating these regions with the processes involved in updating story representations.

In summary, the comparison of grammatically correct sentences with comprehensible versus incomprehensible meanings has been associated with activation in anterior and posterior parts of the left middle temporal gyrus,^{9,19,21} bilateral anterior temporal poles,^{21,22} the left angular gyrus,^{9,19} and the posterior cingulate/precuneus.^{18,19} Each region may play a different role in sentence semantics. For example, seeing concurrent hand gestures enhanced activation for heard sentences in the posterior middle temporal cortex consistent with auditory and visual semantics converging at the posterior end of the semantic network. Inconsistent activation in the left angular gyrus and medial parietal regions may reflect task-dependent semantic retrieval strategies; and activation in anterior middle temporal lobe may reflect the integration of multiple semantic concepts.

When the sentence comprehension activations are illustrated on the cartoon brains in Figure 2, it is striking to see that temporal lobe activation for plausible versus implausible sentences (middle right panel) lies outside the areas associated with prelexical processing (top left panel) or pseudowords (top right panel) but inside the ring of activation associated with semantic processing of single words (middle left panel). This dissociation might reflect a task confound (*e.g.*, lexical decision for words and plausibility judgments for sentences). Alternatively,

it raises the interesting possibility that activation in the superior temporal sulci for sentences with plausible versus implausible meanings reflects the consolidation of multiple semantic concepts into a new and integrated concept. For example, the sentence “My memory is a little cloudy” has six words including three rich content words (memory, little, cloudy) to express one concept “forgotten.” This suggests that activation for accessing the semantics of single words is spatially extensive but consolidates in the ventral banks of the superior temporal sulci when multiple meanings can be integrated into fewer concepts.

Semantic constraints in sentence comprehension

The previous section focused on activation for grammatically correct sentences with plausible versus implausible meanings. In contrast, this section considers the reverse of this comparison, that is, which areas are more activated by sentences with implausible versus plausible meanings. The assumption here is that, when the sentence meaning is implausible, ambiguous, or unconstrained, there will be a greater need for semantic constraints that will be expressed by top-down processing from our prior knowledge of the world and the experimental context.

When sentence comprehension becomes more difficult, activation is consistently observed in the left pars opercularis. I am distinguishing dorsal versus ventral pars opercularis according to whether activation is more or less than 20 mm above the ACPC line. Dorsal pars opercularis activation has been reported for grammatically correct sentences with implausible versus plausible meanings by Ye and Zhou³⁵ at $[-40 22 24]$ and in several studies when the meaning of the sentence is more difficult to extract. For example, by Bilenko and colleagues³⁶ at $[-52 14 23/47 14 29]$ for sentences with ambiguous versus unambiguous meanings; by Mashal and colleagues¹⁹ at $[-44 18 30]$ for sentences with novel metaphoric versus literal meanings; by Willems and colleagues³⁷ at $[-40 10 22]$ when speech was presented with incongruent relative to congruent gestures or pantomimes (*i.e.*, when there was interference at the comprehension level); and by Desai and colleagues³⁸ at $[-54 10 31]$ when participants listened to sentences with abstract versus visual or motoric meanings. However, left dorsal pars opercularis

activation is not specific to semantic contexts. It has also been reported by Myers and colleagues⁵ at $[-41\ 8\ 27]$ for prelexical stimuli when there is an unexpected change in phonetic category and in other studies of linguistic and nonlinguistic sequencing (see next section).

More ventrally, pars opercularis activation has been reported by Tyler and colleagues³⁹ at $[-51\ 15\ 18]$ for grammatically correct sentences with implausible versus plausible meanings; by Obleser and Kotz⁹ at $[-60\ 12\ 16]$ for sentences with low versus high semantic expectancy; by Desai and colleagues³⁸ at $[-51\ 19,\ 13]$ when participants listened to sentences with abstract versus visual or motoric meanings; by Turner and colleagues⁴⁰ at $[-50\ 14\ 18]$ when participants view silent videos of a speaker making speech versus non-speech oral movements; and by Szyck and colleagues⁴¹ at $[-49\ 14\ 18]$ when heard speech is inconsistent with seen oral movements. These activations may reflect top-down processing that is attempting to predict the stimuli. For example, sentences with low semantic expectancy (e.g., “*She weighs the flour*”) are difficult to predict before completion of the sentence because a verb such as “weighed” can be associated with multiple objects; whereas the sentences with high semantic expectancy (e.g., “*She sifts the flour*”) are easier to predict before sentence completion because there are not many nouns that are likely to follow verbs like “sift.” In this context, activation in the left ventral pars opercularis might reflect increased demands on predictions about the forthcoming sequence of events. The examples in this section have pertained to semantic predictions. However, left ventral pars opercularis activation has also been associated with verbal working memory at $[-53\ 7\ 15]$ by Koelsch and colleagues⁴²; and for articulatory planning. Specifically, Papoutsis and colleagues⁴³ report activation at $[-54\ 12\ 12]$ for the repetition and subvocal rehearsal of pseudowords with low- versus high-sublexical frequency; a manipulation that does not involve semantic, syntactic, or phonological word form processing. Left ventral pars opercularis activation can therefore be consistently explained in terms of the generation of semantic or articulatory sequences. Its function is considered again in sections “Articulation” and “Summary of speech production.”

The comparison of auditory sentences with plausible versus implausible meanings was also associ-

ated with left pars orbitalis, activation at $[-39\ 27\ -6/-48\ 24\ -9]$ by Tyler and colleagues³⁹; and at $[-52\ 32\ -4]$ by Ye and Zhou.³⁵ The same region has been reported by Davis and Gaskell¹³ at $[-40\ 27\ -9/-48\ 24\ -12]$ for single auditory words relative to pseudowords; by Nosarti and colleagues⁴⁴ at $[-38\ 32\ -6]$ for reading irregularly spelled words relative to pseudowords; by Schafer and Constable⁴⁵ at $[-42\ 28\ -11]$ for semantic relative to syntactic processing in written speech comprehension; and by Aarts and colleagues⁴⁶ at $[-40\ 22\ -14]$ when the written words “right” or “left” are incongruent with the direction of an arrow (right or left). All these studies suggest that left pars orbitalis activates with increasing demands on semantic retrieval in the context of semantic conflict that may be arising at the single-word level or between the meaning of single words and sentences. Likewise, right inferior frontal activation has also been reported in the context of conflicting semantic information. For example, activation in the right pars opercularis $[44\ 18\ 14/54\ 18\ 12]$ and triangularis $[46\ 28\ 6]$ were reported by Snijders and colleagues²⁶ when the meaning of a sentence was ambiguous versus unambiguous; and by Peelle and colleagues⁴⁷ at $[44\ 14\ 22/44\ 4\ 28]$ when the meanings of a series of sentences conflict with one another. Dick and colleagues³¹ also report right inferior frontal activation when participants listened to and watched a story teller using hand movements that were semantically incongruent relative to congruent with the spoken speech. As with left inferior frontal activation, right inferior frontal activation is not specific to the auditory modality. It has also been reported at $[60\ 28\ -4]$ by Schmidt and Seger⁴⁸ when subjects read metaphors relative to literal sentences. Right inferior frontal activation may therefore increase when there is incongruency between the meaning of the words and the meaning of the sentence.

In summary, when the semantic content of sentences is difficult to extract, activation increases in the left and right pars opercularis and orbitalis. This contrasts to the temporal and parietal regions that are more activated for sentences with plausible meanings. The most likely explanation is that activation in the pars orbitalis reflects semantic competition at the single-word level and activation in the pars opercularis reflects top-down predictions on the plausible sequence of events. This is

discussed in more detail in the following sections that also attempt to clarify a distinction between the function of dorsal and ventral regions of the pars opercularis.

Syntactic constraints

Syntactic processing refers to the hierarchical sequencing of words and their meanings with the expected order of words depending on the language spoken. Consequently, there are language-specific grammatical rules about adjacent words (what type of word is likely to follow another) and long-distance word associations. Deriving the meaning of words becomes particularly complex when one phrase is embedded within another. For example, a sentence such as “the man the dog chased was carrying a ball” has one relative clause (“the dog chased”) nested within a simpler sentence “the man was carrying the ball.” As the sentence becomes more complex, the demands on auditory short-term memory increase because items have to be held in memory for longer until the meaning of the sentence emerges.

Syntactic processing has been investigated by comparing sentences with grammatical errors to sentences without grammatical errors; and for sentences with more versus less syntactically complex structures. In both cases, the demands on syntactic processing are confounded by the differing demands on semantics because both grammatical errors and complex sentences make it more difficult to extract the meaning of the sentence. It is therefore not surprising that the left pars opercularis areas that were more activated for sentences with implausible versus plausible meanings in the previous section are also more activated for sentences with grammatical errors or complex structures. For example, left ventral pars opercularis activation has been reported at [−62 18 12] by Friederici and colleagues¹⁰ when sentences had syntactic errors; and at [−47 12 18] by Raettig and colleagues⁴⁹ when there were violations in verb–argument structure. As discussed in the previous section on semantic constraints, activation in the left ventral pars opercularis has also been associated with verbal working memory [−53 7 15] and predicting the sequence of semantic or articulatory events. It will be discussed again in section “Summary of speech production.”

Syntactic complexity has also been shown to increase auditory sentence activation in the left dorsal

pars opercularis at [−45 6 24], even when working memory demands are factored out.⁵⁰ An effect of syntactic complexity at [−48 16 30] has also been reported by Newman and colleagues⁵¹ for written sentences; and activation at [−46 16 24] during written sentences was more strongly primed by prior presentation of the verbs than the nouns.⁵² Makuuchi and colleagues⁵⁰ attributed dorsal left pars opercularis activation to the hierarchical organization of sequentially occurring events. Hierarchical sequencing of events is a core component of syntactic processing but it is not limited to linguistic stimuli. It can therefore be investigated using nonlinguistic stimuli and such studies have again highlighted the involvement of the left dorsal pars opercularis.^{53,54} For example, Bahlmann and colleagues⁵³ found significantly higher activation in the dorsal pars opercularis at [−50 6 30] when sequences of nonlinguistic visual symbols could be predicted on the basis of nonadjacent word dependencies compared to adjacent word dependencies. Similarly, Tettamanti and colleagues⁵⁴ found dorsal left pars opercularis activation at [−54 12 24] when participants learnt the nonrigid (variable) dependencies of items within a sequence of unfamiliar colored shapes. The dorsal left pars opercularis therefore appears to be involved in sequencing events, irrespective of whether they are linguistic or nonlinguistic.

The third region that has been associated with syntactic errors or complexity lies on the border between the pPT and the vSMG. Specifically, Raettig and colleagues⁴⁹ reported activation at [−65 −42 15] for sentences with grammatical errors (e.g., “Peter has *eat* the apple” vs. “Peter has *eaten* the apple”); Friederici and colleagues⁵⁵ report activation at [−63 −42 21] for syntactically complex versus less complex sentences; and Richardson and colleagues³² report increased activation at [−58 −42 22] for sentences where the meaning depends on the order of the subject and object (e.g., “*The dog chased the horse*”/“*The horse chased the dog*”) versus sentences where the subject and object are not reversible (e.g., “*The dog eats the bone*”).

Activation in the vicinity of pPT and/or vSMG is not limited to difficult syntactic contexts. These areas are also activated when sentence comprehension is made more challenging at the perceptual or semantic level. In perceptually challenging contexts, Dos Santos Sequeira and colleagues⁵⁶ report activation at [−56 −42 24] when participants

listened to consonant-vowel syllables in background noise. In semantically challenging contexts, Oleser and Kotz⁹ report activation at $[-46 -36 18]$ for sentences with low versus high semantic expectancy particularly when the stimuli had low intelligibility; and Hickok and colleagues¹ report pPT/vSMG activation for individual subjects at $[x = -47$ to -61 ; $y = -32$ to -50 ; and $z = 11$ to $28]$ for hearing sentences that were syntactically correct but semantically meaningless because some words in the sentence had been replaced with pseudowords (“*It is the glandor in my nedderop*”).

Why does left pPT/vSMG activate during difficult speech comprehension conditions? Evidence from studies of auditory short-term memory^{42,43,57} and speech production in the absence of auditory stimuli^{32,58} suggests that it is involved in subvocal articulation. For example, Papoutsis and colleagues⁴³ found increased left pPT/vSMG activation bilaterally at $[-56 -38 20/64 -32 10]$ when subjects articulated four versus two syllables during a task that involved delayed repetition and subvocal rehearsal of pseudowords. Left pPT or vSMG activation has also been observed during auditory working memory at $[-44 -38 21]$ by Koelsch and colleagues⁴² and at $[-63 -34 19]$ by Buchsbaum and D’Esposito.⁵⁷ One interpretation is that left pPT/vSMG activation reflects a working memory rehearsal strategy that holds an auditory representation in memory. However, in Buchsbaum and D’Esposito,⁵⁷ the task did not involve the online retention of a fixed set of items. Instead, left pPT/vSMG activation was observed when participants made correct word recognition decisions on constantly changing targets that may or may not have been presented before. Buchsbaum and D’Esposito⁵⁷ therefore suggest that left pPT/vSMG activation is involved in phonological retrieval processes (perhaps subvocal articulation) that occurs automatically during successful recognition of recently presented words.

Other studies have shown that left pPT/vSMG activation is not limited to speech production in the context of auditory stimuli or short-term memory tasks. Nor is it limited to the production of speech sounds. For example, Richardson and colleagues³² report activation at $[-62 -40 14/50 -46 20]$ for alternatively articulating “one” and “three” in response to seeing (not hearing) the digits 1 and 3, and also for nonverbal mouth movements versus hand movements; and Wilson and colleagues⁵⁸ report ac-

tivation at $[-64 -34 24]$ for naming pictures with low- versus high-frequency names. For non-speech sounds, Koelsch and colleagues⁴² report activation at $[-47 -42 24]$ when a sequence of auditory tones is held in memory; and Yoncheva and colleagues⁵⁹ report activation at $[-42 -38 26]$ for a tone matching task (are tone-triplet pairs the same or not) as well as for a rhyme-matching task (do word pairs rhyme or not) with no significant difference in activation for the rhyme versus tone comparisons.

The correspondence in the location of speech and non-speech effects has been demonstrated by Koelsch and colleagues⁴² for auditory working memory tasks; by Yoncheva and colleagues⁵⁹ during matching tasks; and by Richardson and colleagues³² for auditory speech comprehension and visually cued mouth movements. Common activation for two tasks may or may not reflect a common underlying function. Indeed, Hickok and colleagues¹ dissociated the spatial pattern of activation in pPT/vSMG for speech comprehension and production, using multivariate pattern classification. This approach identifies a significant difference in the pattern of voxels over a single region but it does not dissociate the functional contribution of the pPT versus vSMG regions.

To conclude this section, activation for syntactically difficult sentences increases in the left pars opercularis and the left pPT/vSMG.⁵⁵ The same regions are also activated when syntactic demands are held constant and sentence comprehension is made more difficult. Therefore, activation in these regions is not specific to syntactic processing. The dorsal pars opercularis appears to play a top-down role in sequencing linguistic and nonlinguistic events, independent of working memory demands.⁵⁰ The ventral pars opercularis is involved in verbal working memory⁴² and sequencing semantic and articulatory events. In contrast, activation in the pPT and vSMG can be explained by subvocal articulation. The functional distinctions between these regions will be discussed further in section “Summary of speech production” but future studies will be required to dissociate the functions of pPT and vSMG.

Subvocal articulation during speech comprehension

The previous section concluded that left pPT/vSMG activation during difficult speech comprehension reflected subvocal articulation. The close link

between comprehension and articulation is one of the unique features of speech because we can accurately reproduce speech sounds but we cannot accurately reproduce non-speech sounds. This is not surprising given that auditory speech, by definition, is constrained by the auditory signals that our vocal tracts and articulators can produce. Evidence that heard speech activates areas involved in articulating speech comes from common activation for (i) auditory speech in the absence of articulation and (ii) articulation in the absence of auditory speech. Three areas meet these criteria. One is the motor cortex,^{60,61} the others are the pPT and vSMG regions discussed in the previous section. In all three of these regions, activation is consistently activated during speech production but inconsistently activated during speech comprehension. It was quite striking that the majority of auditory speech perception and comprehension studies that I reviewed did not report left pPT or vSMG activation, irrespective of whether the stimuli were sentences,^{10,23,24,38} words,^{8,11,12,62–64} or prelexical consonant vowels.^{3,7} I only found reports of left pPT and/or vSMG activation in five studies of sentence comprehension^{1,9,32,49,55} and one study of prelexical perception.⁵⁶ In all cases, pPT/vSMG activation was identified when speech comprehension was made more difficult at the syntactic, semantic, or perceptual level (see previous section). Thus, fMRI evidence suggests that an area associated with speech production is not consistently involved in speech comprehension (when acoustic inputs are matched) but may play an essential role when speech comprehension is difficult.

A similar conclusion has been reached for the role of the motor cortex in speech perception. This is discussed in two recent review papers.^{60,61} Devlin and Aydelott⁶⁰ discuss the degree to which functional imaging data support the motor theory of speech perception. They emphasize that motor and premotor activation is not consistently reported in speech perception studies but it is observed when speech stimuli are degraded (*e.g.*, embedded in noise) or minimal (*e.g.*, syllables rather than words) and may therefore be recruited to aid speech comprehension, when speech perception is challenging. Scott and colleagues⁶¹ also conclude that motor activation might facilitate speech perception in difficult listening conditions. In addition, they suggest that motor activation during speech perception allows

people to coordinate their speech with others, both in terms of turn taking and also in terms of idiosyncratic characteristics of pronunciation and the use of conceptual and syntactic structures. The main point, however, in both reviews^{60,61} is that activation in motor cortex is not always necessary for phonetic encoding or speech comprehension.

To conclude this section, left motor and left pPT and left vSMG activations have been reported for tasks that challenge auditory comprehension and each of these regions are also activated for speech production in the absence of auditory cues. Activation may therefore reflect subvocal articulation that facilitates difficult speech comprehension. More specifically, the overlap in the location of activation for speech production and difficult speech perception may reflect the use of the speech production system to make predictions during speech perception.

The role of prosody in speech comprehension

Prosody refers to the patterns of stress and intonation in speech. These patterns carry emotional and nonverbal information that supplements the lexico-semantic information carried by the spoken words. Wiethoff and colleagues⁶⁵ found that emotional prosody increases activation in the amygdala. This is consistent with previous studies showing amygdala activation for facial expressions,⁶⁶ and when subjects read emotional words.⁶⁷ Wittfoth and colleagues⁶⁸ found that when the emotional prosody in heard sentences was incongruous with semantic content, activation increased in the right superior temporal gyrus and sulcus and right dorsal anterior cingulate cortex (ACC). Many previous studies have associated the right superior temporal sulcus with nonverbal emotional processing of human voices and faces but Kreifelts and colleagues⁶⁹ have recently demonstrated a functional subdivision of the superior temporal lobes. They found maximum voice sensitivity in the trunk of the superior temporal lobe and maximum face sensitivity in the posterior terminal ascending branch. They also argue that an overlap of these two regions at the bifurcation of the superior temporal cortex may support the formation of an audiovisual emotional percept.

Although the earlier studies have investigated the impact of emotional prosody in general, Ethofer and colleagues⁷⁰ attempted to dissociate activation for pseudowords spoken in five prosodic categories

(anger, sadness, neutral, relief, and joy). Using multivariate pattern analysis, they demonstrated that each emotion had a specific spatial signature in the auditory cortex that generalized across speakers.

Summary of activation related to speech comprehension

In summary, the current speech comprehension literature supports the following anatomical model. Acoustic processing of prelexical speech increases activation in the vicinity of bilateral HG with selectivity for speech emerging in lateral, anterior, ventral, and posterior regions of the superior temporal gyri and sulci. Semantic processing of single words extends even further in the anterior, ventral, and posterior directions into the middle and inferior temporal cortex. Remarkably, however, semantic processing of plausible relative to implausible sentences was more focal than for words, and was observed in the ventral banks of the anterior and posterior superior temporal sulci, which lie between the activations associated with prelexical processing in the superior temporal gyri and single-word semantics in the middle and inferior temporal gyri (see Fig. 2). This constricted pattern of activation for sentence semantics might reflect the consolidation of multiple words into a reduced set of concepts.

The observation that prelexical and semantic processing of spoken words extends in anterior, ventral, and posterior directions suggests that the same speech inputs can follow multiple different pathways. The location of prelexical activation would then be determined by the task demands. For example, when an articulatory response is required, top-down expectations from motor programs may stabilize and enhance prelexical processing in the dorsal processing direction but when a lexical decision is required, top-down expectations from semantic knowledge may stabilize and enhance prelexical activation in the ventral direction. In this way, the cortical networks supporting language comprehension are dynamically determined by the task and context³⁰ (see Fig. 3).

Several different sources of top-down inputs/constraints are distinguished in Figure 3, including semantic, syntactic (word sequences), articulation, auditory, visual (spelling), and spatial location for locomotion. The review reported in this section localized semantic constraints to the left pars orbitalis, ventral pars opercularis, angular gyrus, and

precuneus. Word order constraints (grammatical sequencing) were identified in the dorsal pars opercularis; and articulatory constraints were associated with activation in the left motor cortex,⁶⁰ vSMG and pPT (see Fig. 3). For further discussion of these regions (see section “Summary of speech production”).

Left lateralization for speech comprehension

Differences in the role of the left and right temporal lobes have been of particular interest because lesion studies indicate that it is the left rather than right temporal lobe that is needed for speech recognition and production. In this context it is surprising that the comparison of speech and closely matched non-speech stimuli result in bilateral temporal lobe activation.^{9,10} The specialization and lateralization for temporal lobe speech function may therefore be driven by nonacoustic differences between speech and non-speech stimuli. These nonacoustic differences include our significantly greater experience with speech than non-speech stimuli,¹⁴ the resulting categorical qualities of the perceptual representation of speech,⁷¹ the ability to produce as well as to perceive speech,^{60,61} and the influence of spatial orienting.⁶²

Why should the effect of auditory expertise^{14,23} be greater in the left than right temporal lobe? Here I consider two possibilities: (i) either familiarity with the auditory stimulus increases bottom-up connections from acoustic processing in bilateral superior temporal lobes to representations of semantics, grammar, and articulation in the left frontal cortex; or (ii) familiarity results in more successful top-down predictions from left frontal regions that stabilize acoustic processing in the left temporal lobe more than the right temporal lobe. Adank and Devlin²³ support the latter hypothesis. They suggest that the left-lateralized temporal lobe response they observed when time-compressed speech was recognized might be driven by top-down processing from left-lateralized premotor activation. Kouider and colleagues¹¹ also highlight the influence of top-down inputs on lexical processing to explain why they observed repetition suppression in left HG for words but not pseudowords. Similarly, the left-lateralized anterior temporal lobe response that Leff and colleagues³ observed for vowel deviants could reflect top-down mechanisms (see section “Prelexical phonemic processing”).⁶

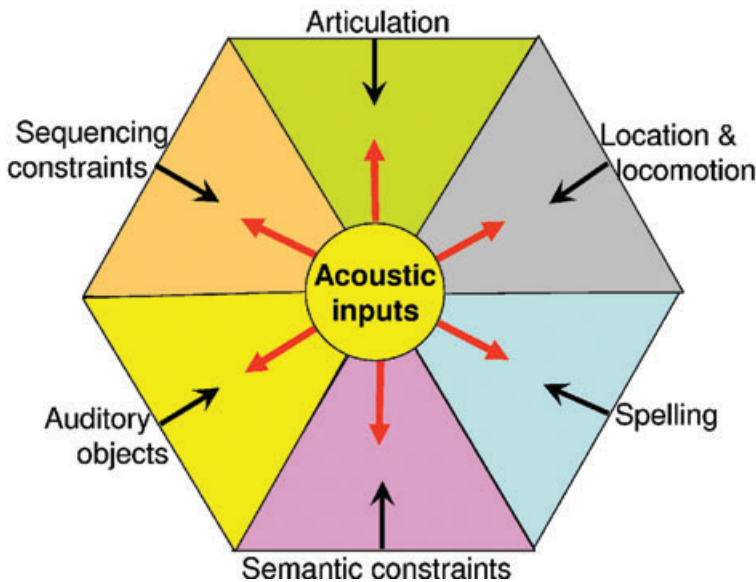
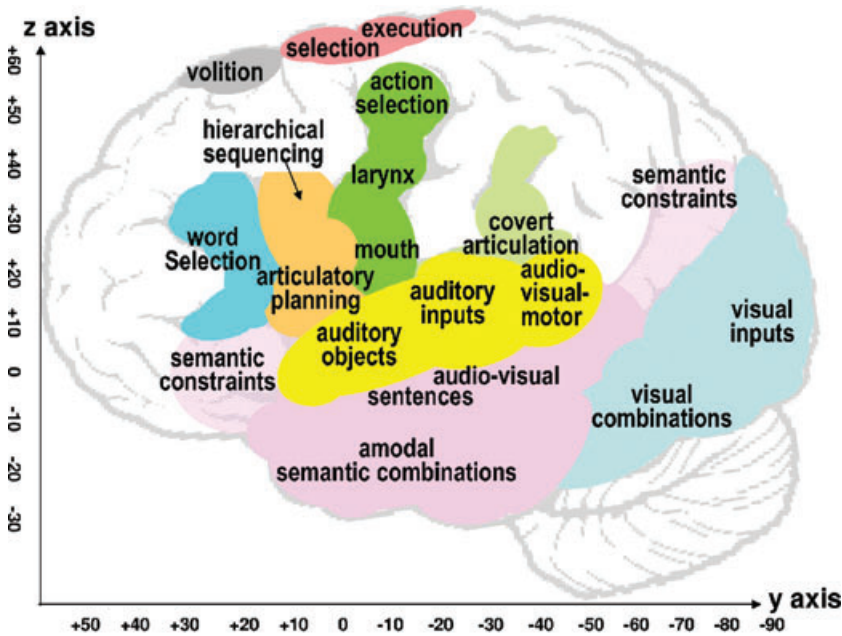


Figure 3. Functional attributions. *Top panel:* Summary of functions attributed to language areas, see text for rationale. Future reviews are required to postulate and illustrate the multiple pathways that connect these areas. *Below panel:* a schematic summary of the balance between bottom-up and top-down processing during speech comprehension. The hypothesis is that top-down constraints from frontal, motor, and parietal areas influence the location of activation foci for bottom-up prelexical processing of auditory inputs.

In conclusion, frontal lobe activations during speech comprehension are more consistently lateralized than temporal lobe activations. The proposal developed here is that lateralization in the temporal lobes may be driven top-down from higher level processing in frontal areas. Moreover, different parts of the frontal cortex may determine lateralization in different temporal regions. If this hypothesis is correct, then the more damaging effect of left temporal lobe lesions compared to right temporal lobe lesions on language function might be a consequence of loss of top-down connections from anterior speech areas. This could be investigated with functional connectivity studies.

Review of speech production

This section focuses on speech production, a complex multistage process that links conceptual ideas to articulation. The cognitive components of speech production have been described in many studies. Its roots are in the conceptual ideas that need to be expressed. The words associated with these concepts must then be retrieved and sequenced with their appropriate morphological forms. At the same time, competition from other words with similar meanings needs to be suppressed. Word selection is therefore a dynamic interaction between excitation of the intended words and inhibition of the unintended words.^{72–74}

Sequencing is required at several different levels of the speech production process. At the sentence level, words need to be sequenced in a grammatically correct order. Within a word, sequencing is required to assemble the correct combination of phonemes and syllables. In addition, pitch, prosody, and a metrical structure need to be applied to speech output,⁷⁵ which will depend on the conversational context. The combination of phonemes and syllables being produced is translated into a sequence of articulatory plans (phonetic encoding). This is followed by the initiation and coordination of sequences of movements in the speech articulators, which include the tongue, lips, and laryngeal muscles that vibrate the vocal tract for vowel phonation and prosody. At the same time, respiration needs to be controlled and the resulting auditory signal needs to be monitored and fed back to the motor system for online correction of speech production.

Given the many processes that support speech production, it is not surprising that many different

brain areas are involved. Each level of the classical speech output pipeline are addressed in the following sections: conceptual processing (controlling for word retrieval and articulation) (section “Conceptual processing in speech production”); word retrieval (controlling for conceptual processing and articulation) (section “Word retrieval”); articulation (controlling for conceptual processing and word retrieval) (section “Articulation”); and auditory feedback (section “Monitoring speech output”).

Conceptual processing in speech production

In a large-scale review of 120 functional neuroimaging studies of the neural systems that store and retrieve semantic memories, Binder and colleagues¹⁶ identified seven left-lateralized regions associated with amodal semantic processing: (i) inferior frontal gyrus, (ii) ventral and dorsal medial prefrontal cortex, (iii) posterior inferior parietal lobe, (iv) middle temporal gyrus, (v) fusiform, (vi) parahippocampal gyri, and (vii) the posterior cingulate gyrus. All these areas have been associated with speech comprehension in the previous section. Their contribution to speech production depends on the task and the type of semantic information that needs to be retrieved.

A demonstration that speech production and speech comprehension access the same conceptual system was provided by de Zubicaray and McMahon.⁷⁶ They found that activation in the left pars orbitalis [$-51\ 24\ -9$] was reduced when picture naming occurred in the context of semantically related heard words versus unrelated or phonologically related heard words. The authors suggest that the picture is recognized before the auditory word and semantic processing of the picture facilitates the recognition of the auditory word. The coordinates of this effect [$-51\ 24\ -9$] also correspond to those reported by Tyler and colleagues³⁹ at [$-48\ 24\ -9$] and by Ye and Zhou³⁵ at [$-52\ 32\ -4$] when heard sentences were semantically implausible versus plausible. The left pars orbitalis is therefore involved in semantic retrieval processes that support both speech comprehension and production tasks.

A direct comparison of semantic versus phonological word retrieval by Birn and colleagues²⁰ has illustrated the role of the medial superior frontal cortex at [$-21\ 11\ 45$] in the conceptual processes that

support speech production. It is interesting to note that activation in the medial superior frontal cortex has also been reported by Tremblay and Gracco⁷⁷ at $[-21\ 19\ 57]$ for “volitional” versus stimulus-driven word selection. This raises the possibility that semantic word retrieval (*e.g.*, producing words that belong to a particular semantic category like “fruits”) requires more volitional effort than phonological word retrieval (*e.g.*, producing words that start with the same letter). The point here is that increased activation for semantic tasks may reflect the demands on executive strategies that are necessary for, but not limited to, semantic word retrieval.

In addition to the widely distributed set of regions that are typically associated with semantic/conceptual processing during speech production and comprehension, many other sensory-motor areas may play a role in accessing words with specific meanings. This was illustrated in a study of verbal fluency by Hwang and colleagues,⁷⁸ who found that the retrieval of words belonging to visual categories activated extrastriate cortex (a secondary visual processing area); retrieval of words belonging to motor categories activated the intraparietal sulcus and posterior middle temporal cortex; and retrieval of words belonging to somato-sensory categories activated postcentral and inferior parietal regions. Thus, they demonstrate that lexico-semantic processing during speech production is distributed across brain regions participating in sensorimotor processing. This is argued to be a consequence of the sensorimotor experiences that occurred during word acquisition.

Hocking and colleagues⁷² also demonstrate how activation during speech production depends on the type of words that are being retrieved. These authors found bilateral hippocampal activation at $[-30\ -3\ -30/33\ -6\ -33]$ when pictures to be named were blocked in terms of their semantic category; and bilateral anterior medial temporal activations when the objects to be named were blocked according to similar visual features. Whitney and colleagues⁷⁹ also found left hippocampal activation $[-28\ -35\ 5]$ for spontaneous word production (free verbal association) relative to category and letter fluency. These medial temporal lobe activations during speech production tasks therefore appear to depend on the demands placed on semantic retrieval even though medial temporal lobe structures such as the hippocampus are more classically associated with the

initial acquisition of new words.⁸⁰ To conclude, conceptual processing during speech production activates the same set of regions that have been associated with single-word comprehension.

Word retrieval

To investigate the processing that supports word retrieval, speech production has been compared to tasks that control for articulation. For example, Whitney and colleagues⁷⁹ and Jeon and colleagues⁸¹ compared word generation tasks (find, select, and produce a word related to the stimulus) to reading (produce the word or nonword specified by the stimulus). In both studies, the most significant and consistent effects were observed in the left inferior and middle frontal gyri, spanning both the pars opercularis (BA 44), the pars triangularis (BA45), and the inferior frontal sulcus. The peak coordinates for the left middle frontal activations in these two studies are remarkably similar: at $[-48\ 28\ 21]$ in Whitney and colleagues⁷⁹ and at $[-51\ 25\ 25]$ in Jeon and colleagues⁸¹. Two other studies also demonstrated that this left middle frontal area is involved in word retrieval irrespective of the type of words that needed to be produced. Specifically, Heim and colleagues⁸² observed that the same left middle frontal region at $[-44\ 28\ 22]$ was activated for generating words that were either semantically or phonologically related to a word; and de Zubicaray and McMahon⁷⁶ found an effect of semantic and phonological priming on picture naming at $[-48\ 21\ 18]$, which is again consistent with a generic role for this region in word retrieval.

A second region that is consistently activated in word retrieval tasks is the left dorsal pars opercularis that was associated with top-down sequencing of linguistic and nonlinguistic events in the speech comprehension section (“Syntactic constraints”). During speech production, left dorsal pars opercularis activation was located at $[-40\ 17\ 25]$ by Jeon and colleagues⁸¹ for word generation more than reading; and at $[-49\ 13\ 29]$ by Fridriksson and colleagues⁸³ for imitating or observing a speaker producing nonsense syllables. As viewing a speaker producing nonsense syllables does not involve semantics or articulation,⁸³ speech production activation in the left dorsal pars opercularis is not specific to semantic or syntactic processing. Indeed, as discussed in the speech comprehension section earlier, the left dorsal pars opercularis is involved in the

hierarchical sequencing of linguistic and nonlinguistic events.^{37,50,53,54} This is required for speech production as well as speech comprehension. The left dorsal pars opercularis may therefore play a role in the hierarchical sequencing of events during production as well as comprehension tasks.

A third region that is consistently activated in word retrieval tasks is the left ventral pars opercularis that was associated with predicting semantic or articulatory sequences in the speech comprehension section (“Syntactic constraints”). During speech production, this region has been reported at [−57 9 9] by Zheng and colleagues⁸⁴ for articulating versus listening to the word “Ted”; and at [−54 12 12] by Papoutsi and colleagues⁴³ for the repetition and subvocal rehearsal of pseudowords with low versus high sublexical frequency. As this manipulation does not involve semantic, syntactic, or phonological word form processing, the authors attribute their activations to phonetic encoding (articulatory planning). Left ventral pars opercularis during speech production has also been reported in two studies by Heim and colleagues.^{82,85} In one study, Heim and colleagues⁸² found that left ventral pars opercularis activation at [−46 10 4] was greater for retrieving words that were phonologically versus semantically related to a target. Another⁸⁵ found left ventral pars opercularis activation at [−50 6 6] during picture naming in German was sensitive to the gender of the noun being produced. This was probed by presenting pictures to be named with an auditory word and its gender determiner (e.g., “der Tisch”—the table). The gender of the auditory word and picture were either the same or different with more activation when they were different. However, it is not clear from this study whether activation in the left ventral pars opercularis was sensitive to the gender of the determiner (which would reflect syntactic processing) or to the phonological overlap in the determiner being heard and the determiner being retrieved (which could reflect articulatory processes). The explanation that is most consistent with all the other data being presented is that left ventral pars opercularis activation reflects the demands on articulatory planning.

The distinction between ventral and dorsal pars opercularis has been rather arbitrary. I have been using “dorsal” to refer to activation that is 20–30 mm above the ACPC line and “ventral” when activation is 4–20 mm above the ACPC line but there may be

multiple functional subregions within each of these regions. Moreover, some studies are at the border of ventral and dorsal pars opercularis activation. For example, Heim and colleagues⁸⁶ report activation at [−50 8 19] when pictures to be named were blocked according to similarity in semantics or syntax (gender) versus phonology. The results were interpreted in terms of decreased demands when name retrieval was primed by phonologically related responses and increased demands when stimuli are presented with semantic and syntactic competitors. However, without further studies it is not possible to determine whether the function of this mid-pars opercularis region corresponds to the properties of the dorsal or ventral regions.

In summary, word retrieval activates the left middle frontal cortex and dorsal and ventral left pars opercularis. Left middle frontal activation was not associated with speech comprehension, unless the task required explicit semantic comparisons between two words,¹⁶ therefore the left middle frontal cortex appears to be more involved in word retrieval than word comprehension. In contrast, activation in the dorsal and ventral regions of the pars opercularis has been reported for both speech production and comprehension (see Fig. 2). The dorsal pars opercularis has been associated with sequencing linguistic and nonlinguistic events and the ventral pars opercularis has been associated with sequencing articulatory events.⁴³ Generating or interpreting the sequence of events is essential for speech comprehension and production at both the individual word (i.e., the sequence of phonemes) and sentence level (i.e., the sequence of words). The dorsal and ventral pars opercularis may contribute to sequencing by top-down constraints from prior knowledge of what event typically follows another.

Articulation

When speech is produced, activation typically increases bilaterally in motor and premotor cortex, the cerebellum, the supplementary motor area (SMA and pre-SMA), the superior temporal gyri, the temporo-parietal cortices (PPT/vSMG), and the anterior insula, with left-lateralized activation in the putamen (see Brown *et al.*,⁸⁷ for a recent review). None of these regions are dedicated to speech *per se* as illustrated by Chang and colleagues⁸⁸ who report the same areas in the production of speech and the production of non-speech sounds from

orofacial and vocal tract gestures (e.g., cough, sigh, kiss, snort, laugh, tongue click, whistle, cry) that have no phonemic content. The only areas that Chang and colleagues⁸⁸ found to activate for speech more than vocal tract gestures were the ACC and bilateral caudate. Both these areas have been associated with suppression of inappropriate responses by Aarts and colleagues,⁴⁶ Ali and colleagues,⁷⁴ and Kircher and colleagues.⁸⁹ Greater activation for speech than non-speech production may therefore reflect the greater demands on response selection in the context of more competitors that need to be suppressed during the selection of speech versus non-speech movements.

The functions of the anterior cingulate and caudate nuclei during word selection are likely to be different. There are many subregions in the ACC that may each have their own function during speech production.^{74,90,91} The head of the caudate has also been implicated in sequencing temporal and cognitive processes,⁹² including the sequencing of nonlinguistic events.⁵³ It is therefore plausible that greater activation in the head of the caudate nuclei for speech relative to non-speech⁸⁸ might contribute to greater precision in the sequencing of speech relative to non-speech sounds.

Other studies have shown that activation in areas common to speech and non-speech vocal production (*i.e.*, the premotor, parietal, temporal, cerebellar, and putamen) increases with the length of the spoken utterance⁴³ but is inversely related to the familiarity of the stimulus. For example, Papoutsis and colleagues⁴³ found more activation throughout the speech production system for the repetition of pseudowords with four syllables versus two syllables; and Shuster⁹³ found activation in this system increased for repetition of pseudowords compared to the repetition of words. Interestingly, the most extensive difference between pseudoword and word repetition in Shuster⁹³ was located in the left anterior insula. This is consistent with a study by Moser and colleagues,⁹⁴ who found increased activation in the left anterior insula for repetition of pseudowords with novel syllables relative to the repetition of pseudowords with native (*i.e.*, familiar) syllables.

With respect to the function of the left anterior insula, the meta-analysis by Brown and colleagues⁸⁷ found that the frontal operculum and medial-adjacent anterior insula were activated by syllable singing as well as oral reading; and Koelsch

and colleagues⁴² found left anterior insula activation for the rehearsal of tone (pitch) information [$-32\ 19\ 3$] as well as verbal information [$-32\ 15\ 4$]. Brown and colleagues⁸⁷ speculate that the anterior insula is involved in generalized orofacial functions, including lip movement, tongue movement, and vocalization. The fact that activation is not dependent on whether speech is overt or covert;⁸³ and not reported to depend on the number syllables being produced⁴³ is consistent with previous claims that the anterior insula is involved in the planning rather than execution of articulation. In this context, increased activation in the left anterior insula for unfamiliar speech sounds^{93,94} may simply reflect greater demands on articulatory speech plans when they are unfamiliar versus familiar.

With respect to the function of the premotor cortex, Brown and colleagues⁸⁷ distinguish areas that control larynx movements from those that control tongue and mouth movements. In Talairach coordinates, Brown and colleagues⁸⁷ locate the larynx motor cortex to [$-40\ -12\ 30/44\ -10\ 30$] and locate tongue movements to part of the rolandic operculum at [$-50\ -9\ 23/59\ -5\ 17$], consistent with other studies of non-speech oral movements.⁹⁵ In contrast, Meister and colleagues⁹⁶ found that the most dorsal part of the premotor cortex [$-48\ -12\ 54/\ -45\ -12\ 60$], above the larynx motor area, is activated by finger tapping as well as articulation. This led Meister and colleagues⁹⁶ to suggest that the dorsal premotor region plays a role in action selection and planning within the context of arbitrary stimulus–response mapping tasks. Thus, there are at least three functionally distinct regions in the premotor cortex. Of these three regions, Brown and colleagues⁸⁷ note that activation during speech is typically strongest in areas associated with phonation (*i.e.*, the vibration of the vocal tracts). This is consistent with the findings of Fridriksson and colleagues,⁸³ who show peak activation for overt articulation relative to observation of articulation at [$-58\ -2\ 36/58\ -10\ 34$], which is in the vicinity of the larynx area rather than the mouth and tongue area.

The direct comparison of producing versus observing speech production⁸³ identified the pre-SMA [$2\ 6\ 60$] and left putamen [$-24\ -6\ 6$] as well as the premotor cortex discussed earlier. This suggests that these regions are involved in the initiation or execution of speech movements but activation is

not specific to speech movements. For example, Tremblay and Gracco⁷⁷ observed pre-SMA activation at [−6 13 50] for both volitional and stimulus driven mouth movements, irrespective of whether the response was speech or oral motor gestures; and Bahlmann and colleagues⁵³ report activation at [1 4 57] for learning the hierarchical organization of sequentially occurring nonlinguistic visual symbols.

In summary, articulatory planning of orofacial movements activates the left anterior insula.⁸⁷ This area is equally activated for overt articulation and action observation⁸³; and more activated when the motor plans are unfamiliar.^{93,94} The initiation and execution of movement increases activation in bilateral premotor/motor cortex, the pre-SMA, and the left putamen.⁸³ At least three distinct areas in the premotor/motor cortex support speech production. The most ventral is involved in tongue movements, a more dorsal region is involved in control of the larynx, and the most dorsal region is involved in planning both speech and finger tapping movements. Two areas were found to be more involved with articulation of speech than non-speech orofacial movements. These were the anterior cingulate and bilateral head of caudate. Both regions have been associated with the suppression of competing responses (speech and non-speech), which suggests that the subtleties and precision involved in correct speech production requires more suppression of competition from nontargets (*i.e.*, all the words we know with similar articulatory patterns). The next section discusses the role of the superior temporal gyri, pPT, vSMG, and cerebellum in speech production.

Monitoring speech output

The final stage of speech production involves auditory and somato-sensory monitoring of the spoken response. This is crucial for online correction of speech production, for example, when speakers modify the intensity of their speech in noisy environments, or when auditory feedback is altered (*e.g.*, by delay on the telephone). There are three levels at which this feedback can occur. One is from hearing the sound of the spoken response (auditory feedback), the second is higher level “phonological” processing, and the third is from somato-sensory information derived from the speech movements.^{72,97}

Auditory processing of the spoken response activates bilateral superior temporal regions associ-

ated with speech perception. This has been observed when auditory processing of the spoken response was not controlled with a speech output baseline, for example, by Heim and colleagues⁸⁶ and Wilson and colleagues⁵⁸ during picture naming and by Brown and colleagues⁸⁷ during reading aloud. When speech output is controlled, however, superior temporal activation has not been reported for word generation tasks.^{79,81}

With respect to higher level phonological processing of the spoken output, three recent studies have highlighted the involvement of the pPT and/or vSMG, when speech production is made more difficult and therefore requires more auditory monitoring. In Hocking *et al.*,⁷² activation at [−54 −36 15] increased when pictures were named in the context of semantic interference; and Abel and colleagues,⁷³ showed that picture naming activation at [−56 −37 15] was stronger for phonological than semantic interference. The argument here is that both semantic and phonological interference make picture naming more error prone and pPT and/or vSMG activation is involved in monitoring the verbal response to reduce errors. The effect of verbal self-monitoring on left pPT/vSMG activation was directly demonstrated by Zheng and colleagues⁸⁴ Their participants either whispered or produced the word “Ted.” During production, the auditory input was either (i) consistent with the speech output (*i.e.*, the word “Ted”) or (ii) a masking noise that prevented participants hearing the sound of their own speech. Left pPT/vSMG activation at [−66 −45 15] was greatest when participants spoke in the context of a mask relative to the other conditions. This effect could not be explained by either speaking or hearing *per se* but indicates greater pPT/vSMG activation when there was conflict between the expected auditory input and the actual auditory input. This is consistent with the interaction between auditory inputs and auditory predictions that are generated by the speech production system or conversely motor efferents and motor predictions that are generated from the auditory speech inputs, Zheng and colleagues.⁸⁴

Linking back to the speech comprehension section, pPT/vSMG activation was attributed to subvocal articulation in perceptually or semantically challenging contexts. Increased subvocal articulation can also explain increased pPT/vSMG activation during speech production in the context of semantic,⁷² phonological,⁷³ or acoustic

interference.⁸⁴ Specifically, pPT/vSMG activation could reflect increased subvocal articulation from semantic and phonological competitors^{72,73} that is then suppressed before motor output. There may also be increased subvocal articulation during production when no corresponding articulatory signals are received from auditory processing of the output.⁸⁴ In both cases, pPT and/or vSMG may be involved in suppressing articulatory errors.

With respect to somato-sensory feedback, in the absence of auditory feedback, Peschke and colleagues⁹⁷ suggest a role for the postcentral gyrus. Consistent with this proposal, Zheng and colleagues⁸⁴ report activation in the left postcentral gyrus when auditory feedback during speech production was masked; and Hocking and colleagues⁷² report activation in the right postcentral gyrus during error-prone speech production. It is also relevant to note that, during the same conditions that evoked postcentral activation, both Zheng and colleagues⁸⁴ and Hocking and colleagues⁷² also report activation in the left cerebellum. This would be consistent with a role for both the postcentral gyri and the left cerebellum in somato-sensory feedback during speech production.

In summary, auditory monitoring and feedback during speech production requires the integration of auditory, articulatory, and somato-sensory signals with the motor output. The dynamics of this system have been discussed by van de Ven and colleagues,⁹⁸ who used spatial and temporal independent components analysis to show the dynamic coupling between different functional networks and thereby demonstrated a negative correlation between activation in the bilateral auditory cortex and the supplementary motor area [$-1 -10 58$].

Summary of speech production

The first thing to point out is that many of the areas associated with speech production were also discussed in the previous section on speech comprehension. The overlap in areas associated with conceptual processing is not surprising but the overlap in ventral and dorsal regions of the pars opercularis, pPT, and vSMG requires more consideration. Activation in both the dorsal and ventral pars opercularis has been associated with top-down expectations on the predicted sequence of events, which is required at multiple different stages during both speech comprehension and speech production. The

dorsal pars opercularis is involved in sequencing linguistic and nonlinguistic events and appears to be independent of the demands on working memory. In contrast, the ventral pars opercularis is involved in sequencing articulatory events and is involved in working memory. It is therefore possible that increased ventral pars opercularis activation during difficult speech comprehension^{9,10,38} reflects top-down predictions about the possible or expected articulatory codes associated with the heard words.

Activation in the pPT/vSMG regions has also been associated with subvocal articulation but the response properties in pPT/vSMG and the left ventral pars opercularis are distinct.^{42,49,84} For example, Raettig and colleagues⁴⁹ reported ventral pars opercularis activation for errors in verb–noun agreement (“*Peter has laughed the apple*”) but pPT/vSMG activation for morphosyntactic errors (“*Peter has eat the apple*”). One simplistic interpretation of this is that morphosyntactic errors and pPT/vSMG activation reflect a mismatch in the predicted articulation at the level of a single event. In contrast, verb–noun agreement errors and ventral pars opercularis activation reflect a mismatch in the predicted articulation at the level of sequences of articulatory events. This might also explain why Koelsch and colleagues⁴² reported ventral pars opercularis activation for memorizing lists of heard syllables at the same time as singing, a combination of tasks that would provide conflicting predictions about the forthcoming sequence of events.

Areas that are more specific to speech production were the left middle frontal gyrus, left anterior insula, left putamen, bilateral head of caudate, anterior cingulate, pre-SMA, SMA, motor cortex, and cerebellum. The left middle frontal gyrus has been associated with word retrieval when articulation is controlled; the left anterior insula been associated with articulatory planning irrespective of whether the sound is produced or not; and the left putamen, pre-SMA, and motor cortex have been associated with the initiation and execution of overt relative to covert speech.

Despite the new insights into the neural processing of speech production, the number of functional imaging studies of speech production has been less than the number of speech production studies. This is likely to be a consequence of speech production being more difficult to study than speech comprehension. The difficulties studying speech

production arise because it involves mouth movements that can cause artifacts in the imaging signal as well as increasing head movement. In previous years, this led to a preference for speech production paradigms that used covert rather than overt responses. New techniques for minimizing mouth movements and measuring overt responses in the scanner have helped to overcome this previous limitation.⁴³ There have also been considerable challenges in studying speech production at the narrative rather than single-word level. This is because, during narrative speech production, it is very difficult to constrain and control the speech production rate, the order of words that are retrieved and the choice of words relating to a semantic theme.

As the experimental and technical challenges associated with studies of speech production are conquered, many new questions can be addressed. For example, there has been remarkably little discussion of the cerebellum in any of the papers I reviewed, even though it is clearly activated during articulation,⁸⁷ working memory,⁴² word acquisition,⁸⁰ and auditory self-monitoring during speech production.^{72,84} Eickhoff and colleagues⁹⁹ have also suggested that, during verbal fluency, activity in the cerebellum is fed to the left anterior insula. There has also been little discussion in the reviewed papers on the degree to which activation is lateralized during speech production. My overall impression is one where motor, premotor, subcortical, and superior temporal activation are bilaterally activated but middle frontal activation is left lateralized. In contrast, Birn and colleagues²⁰ reported that temporo-parietal activation was right-lateralized for automatic speech production (generating the names of the months) compared to semantic (category) and phonological (letter) speech production. Further studies are clearly required to understand the determinants of lateralization and this is going to require the estimation of lateralization on a region by region basis.¹⁰⁰

Future directions

This review has considered the functional anatomy of language but it has not drawn any conclusions about how the different regions are functionally connected or how the systems vary across individuals. Many studies have started to investigate the functional connectivity of the language sys-

tem either by exploring the anatomical connections between regions,¹⁰¹ conducting functional connectivity analyses,^{82,98,99,102–107} or by combining fMRI data with neurophysiological techniques that have high temporal resolution including MEG,^{108,109} ERP,^{110–112} and TMS.^{113–115} There is also a growing trend to use anatomically constrained computational modeling.^{116–118} Inter-subject variability in the language networks is also being explored in a number of ways: for example, on the basis of behavioral variations,⁵⁸ age,^{39,63,64,119} multilingualism,^{119–129} lesion studies,¹³⁰ or using unbiased classification schemes.^{131,132} There is also increasing evidence that individual differences in activation are related to the underlying brain structure¹⁰⁰ and that brain structure predicts language performance.^{128,129,133,134} Therefore, future investigations can be embellished by the combination of structural, functional, and behavioral data.^{39,64}

In addition to establishing the anatomical and functional connectivity of the language system, many more studies are still required to increase the spatial resolution of the anatomical conclusions that are currently being drawn. For example, future studies are required to distinguish the contribution of different subregions in the pars opercularis.¹⁰⁷ One relatively novel approach is to examine how the pattern of activity within a specified area (rather than the amplitude within an area) varies with the stimuli and task.^{1,70,71,119} Speech selectivity can also be examined by cross-referencing between auditory activation in animals and humans.¹³⁵ However, as the cognitive function of any region depends on the areas that it interacts with, it may only be possible to dissociate the functions of speech regions by examining how they interact with other brain regions; and how these regional interactions are modulated by task demands.

Acknowledgments

This work was funded by the Wellcome Trust. The author thanks Joseph Devlin, Sue Ramsden, Fiona Richardson, Alex Leff, Mohamed Seghier, David Green, Goulven Josse, and Tom Schofield for their helpful comments.

Conflicts of interest

The author declares no conflicts of interest.

References

1. Hickok, G., K. Okada & J.T. Serences. 2009. Area spt in the human planum temporale supports sensory-motor integration for speech processing. *J. Neurophysiol.* **101**: 2725–2732.
2. Rauschecker, J.P. & S.K. Scott. 2009. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* **12**: 718–724.
3. Leff, A.P., P. Iverson, T.M. Schofield, *et al.* 2009. Vowel-specific mismatch responses in the anterior superior temporal gyrus: an fMRI study. *Cortex* **45**: 517–526.
4. Britton, B., S.E. Blumstein, E.B. Myers & C. Grindrod. 2009. The role of spectral and durational properties on hemispheric asymmetries in vowel perception. *Neuropsychologia* **47**: 1096–1106.
5. Myers, E.B., S.E. Blumstein, E. Walsh & J. Eliassen. 2009. Inferior frontal regions underlie the perception of phonetic category invariance. *Psychol. Sci.* **20**: 895–903
6. Schofield, T.M., P. Iverson, S.J. Kiebel, *et al.* 2009. Changing meaning causes coupling changes within higher levels of the cortical hierarchy. *Proc. Natl. Acad. Sci. USA* **106**: 11765–11770.
7. Specht, K., B. Osnes & K. Hugdahl. 2009. Detection of differential speech-specific processes in the temporal lobe using fMRI and a dynamic “sound morphing” technique. *Hum. Brain Mapp.* **30**: 3436–3444.
8. Vaden, K.I. Jr., L.T. Muftuler & G. Hickok. 2009. Phonological repetition-suppression in bilateral superior temporal sulci. *Neuroimage* **49**: 1018–1023.
9. Obleser, J. & S.A. Kotz. 2009. Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb. Cortex*. doi: 10.1093/cercor/bhp128.
10. Friederici, A.D., S.A. Kotz, S.K. Scott & J. Obleser. 2009. Disentangling syntax and intelligibility in auditory language comprehension. *Hum. Brain Mapp.* doi: 10.1002/hbm.20878.
11. Kouider, S., V. de Gardelle, S. Dehaene, *et al.* 2009. Cerebral bases of subliminal speech priming. *Neuroimage* **49**: 922–929.
12. Kotz, S.A., A. D’Ausilio, T. Raettig, *et al.* 2010. Lexicality drives audio-motor transformations in Broca’s area. *Brain Lang.* **112**: 3–11.
13. Davis, M.H. & M.G. Gaskell. 2009. A complementary systems account of word learning: neural and behavioural evidence. *Phil. Trans. R. Soc. B.* **364**: 3773–3800.
14. Leech, R., L.L. Holt, J.T. Devlin & F. Dick. 2009. Experience with artificial nonspeech sounds recruits speech-sensitive cortical regions. *J. Neurosci.* **29**: 5234–5239.
15. Sharp, D.J., J.E. Warren, M. Awad, *et al.* 2009. The neural response to changing semantic and perceptual complexity during language processing. *Hum. Brain Mapp.* doi: 10.1002/hbm.20871.
16. Binder, J.R., R.H. Desai, W.W. Graves & L.L. Conant. 2009. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex.* **19**: 2767–2796.
17. Diaz, M.T. & G. McCarthy. 2009. A comparison of brain activity evoked by single content and function words: An fMRI investigation of implicit word processing. *Brain Res.* **1282**: 38–49.
18. Whitney, C., W. Huber, J. Klann, *et al.* 2009. Neural correlates of narrative shifts during auditory story comprehension. *Neuroimage* **47**: 360–366.
19. Mashal, N., M. Faust, T. Hendler & M. Jung-Beeman. 2009. An fMRI study of processing novel metaphoric sentences. *Laterality* **14**: 30–54.
20. Birn, R.M., L. Kenworthy, L. Case, *et al.* 2009. Neural systems supporting lexical search guided by letter and semantic category cues: a self-paced overt response fMRI study of verbal fluency. *Neuroimage* **49**: 1099–1107.
21. Rogalsky, C. & G. Hickok. 2009. Selective attention to semantic and syntactic features modulates sentence processing networks in anterior temporal cortex. *Cereb. Cortex* **19**: 786–796.
22. Visser, M., E. Jefferies & M.A. Lambon Ralph. 2009. Semantic processing in the anterior temporal lobes: a meta-analysis of the functional neuroimaging literature. *J. Cogn. Neurosci.* doi: 10.1162/jocn.2009.21309.
23. Adank P. & J.T. Devlin. 2009. On-line plasticity in spoken sentence comprehension: adapting to time-compressed speech. *Neuroimage* **49**: 1124–1132.
24. Devauchelle, A.D., C. Oppenheim, L. Rizzi, *et al.* 2009. Sentence syntax and content in the human temporal lobe: an fMRI adaptation study in auditory and visual modalities. *J. Cogn. Neurosci.* **21**: 1000–1112.
25. Hubbard, A.L., S.M. Wilson, D.E. Callan & M. Dapretto. 2009. Giving speech a hand: gesture modulates activity in auditory cortex during speech perception. *Hum. Brain Mapp.* **30**: 1028–1037.
26. Snijders, T.M., T. Vosse, G. Kempen, *et al.* 2009. Retrieval and unification of syntactic structure in sentence comprehension: an FMRI study using word-category ambiguity. *Cereb. Cortex* **19**: 1493–1503.
27. Kircher, T., B. Straube, D. Leube, *et al.* Neural interaction of speech and gesture: differential activations

- of metaphoric co-verbal gestures. *Neuropsychologia* **47**: 169–179.
28. Holle, H., J. Obleser, S.A. Rueschemeyer & T.C. Gunter. 2009. Integration of iconic gestures and speech in left superior temporal areas boosts speech comprehension under adverse listening conditions. *Neuroimage* **49**: 875–884.
 29. Robins, D.L., E. Hunyadi & R.T. Schultz. 2009. Superior temporal activation in response to dynamic audiovisual emotional cues. *Brain Cogn.* **69**: 269–278.
 30. Skipper, J.I., S. Goldin-Meadow, H.C. Nusbaum & S.L. Small. 2009. Gestures orchestrate brain networks for language understanding. *Curr. Biol.* **19**: 661–667.
 31. Dick, A.S., S. Goldin-Meadow, U. Hasson, *et al.* 2009. Co-speech gestures influence neural activity in brain regions associated with processing semantic information. *Hum. Brain Mapp.* **30**: 3509–3526.
 32. Richardson, F.M., M.S. Thomas & C.J. Price. 2009. Neuronal activation for semantically reversible sentences. *J. Cogn. Neurosci.* doi: 10.1162/jocn.2009.21277.
 33. Carreiras, M., M.L. Seghier, S. Baquero, *et al.* 2009. An anatomical signature for literacy. *Nature* **461**: 983–986.
 34. Brownsett, S.L. & R.J. Wise. 2009. The contribution of the parietal lobes to speaking and writing. *Cereb. Cortex.* doi: 10.1093/cercor/bhp120.
 35. Ye, Z. & X. Zhou. 2009. Conflict control during sentence comprehension: fMRI evidence. *Neuroimage* **48**: 280–290.
 36. Bilenko, N.Y., C.M. Grindrod, E.B. Myers & S.E. Blumstein. 2009. Neural correlates of semantic competition during processing of ambiguous words. *J. Cogn. Neurosci.* **21**: 960–975.
 37. Willems, R.M., A. Ozyürek & P. Hagoort. 2009. Differential roles for left inferior frontal and superior temporal cortex in multimodal integration of action and language. *Neuroimage* **47**: 1992–2004.
 38. Desai, R.H., J.R. Binder, L.L. Conant & M.S. Seidenberg. 2010. Activation of sensory-motor areas in sentence comprehension. *Cereb. Cortex.* **20**: 468–478.
 39. Tyler, L.K., M.A. Shafto, B. Randall, *et al.* 2010. Preserving syntactic processing across the adult life span: the modulation of the frontotemporal language system in the context of age-related atrophy. *Cereb. Cortex.* **20**: 352–364.
 40. Turner, T.H., J. Fridriksson, J. Baker, *et al.* 2009. Obligatory Broca's area modulation associated with passive speech perception. *Neuroreport* **20**: 492–496.
 41. Szyck, G.R., H. Jansma & T.F. Münte. 2009. Audiovisual integration during speech comprehension: an fMRI study comparing ROI-based and whole brain analyses. *Hum. Brain Mapp.* **30**: 1990–1999.
 42. Koelsch, S., K. Schulze, D. Sammler, *et al.* 2009. Functional architecture of verbal and tonal working memory: an fMRI study. *Hum. Brain Mapp.* **30**: 859–873.
 43. Papoutsis, M., J.A. de Zwart, J.M. Jansma, *et al.* 2009. From phonemes to articulatory codes: an fMRI study of the role of Broca's area in speech production. *Cereb. Cortex* **19**: 2156–2165.
 44. Nosarti, C., A. Mechelli, D.W. Green & C.J. Price. 2010. The impact of second language learning on semantic and nonsemantic first language reading. *Cereb. Cortex* **20**: 315–327.
 45. Schafer, R.J. & T. Constable. 2009. Modulation of functional connectivity with the syntactic and semantic demands of a Noun Phrase Formation Task: a possible role for the Default Network. *Neuroimage* **46**: 882–890.
 46. Aarts, E., A. Roelofs & M. van Turenout. 2009. Attentional control of task and response in lateral and medial frontal cortex: brain activity and reaction time distributions. *Neuropsychologia* **47**: 2089–2099.
 47. Peelle, J.E., V. Troiani & M. Grossman. 2009. Interaction between process and content in semantic memory: an fMRI study of noun feature knowledge. *Neuropsychologia* **47**: 995–1003.
 48. Schmidt, G.L. & C.A. Seger. 2009. Neural correlates of metaphor processing: the roles of figurativeness, familiarity and difficulty. *Brain Cogn.* **71**: 375–386.
 49. Raettig, T., S. Frisch, A.D. Friederici & S.A. Kotz. 2009. Neural correlates of morphosyntactic and verb-argument structure processing: an fMRI study. *Cortex.* doi: 10.1016/j.cortex.2009.06.003.
 50. Makuuchi, M., J. Bahlmann, A. Anwender & A.D. Friederici. 2009. Segregating the core computational faculty of human language from working memory. *Proc. Natl. Acad. Sci. USA* **106**: 8362–8367.
 51. Newman, S.D., D. Lee & K.L. Ratliff. 2009. Off-line sentence processing: what is involved in answering a comprehension probe? *Hum. Brain Mapp.* **30**: 2499–2511.
 52. Newman, S.D., K. Ratliff, T. Muratore & T. Burns Jr. 2009. The effect of lexical priming on sentence comprehension: an fMRI study. *Brain Res.* **1285**: 99–108.
 53. Bahlmann, J., R.I. Schubotz, J.L. Mueller, *et al.* 2009. Neural circuits of hierarchical visuo-spatial sequence processing. *Brain Res.* **1298**: 161–170.
 54. Tettamanti, M., I. Rotondi, D. Perani, *et al.* 2009. Syntax without language: neurobiological evidence for

- cross-domain syntactic computations. *Cortex* **45**: 825–838.
55. Friederici, A.D., M. Makuuchi & J. Bahlmann. 2009. The role of the posterior superior temporal cortex in sentence comprehension. *Neuroreport* **20**: 563–568.
 56. Dos Santos Sequeira, S., K. Specht, M. Moosmann, *et al.* 2009. The effects of background noise on dichotic listening to consonant-vowel syllables: an fMRI study. *Laterality*. doi: 10.1080/13576500903045082.
 57. Buchsbaum, B.R. & M. D'Esposito. 2009. Repetition suppression and reactivation in auditory-verbal short-term recognition memory. *Cereb. Cortex*. **19**: 1474–1485.
 58. Wilson, S.M., A.L. Isenberg & G. Hickok. 2009. Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. *Hum. Brain Mapp.* **30**: 3596–3608.
 59. Yoncheva, Y.N., J.D. Zevin, U. Maurer & B.D. McCandliss. 2009. Auditory selective attention to speech modulates activity in the visual word form area. *Cereb. Cortex*. doi:10.1093/cercor/bhp129.
 60. Devlin, J.T. & J. Aydelott. 2009. Speech perception: motoric contributions versus the motor theory. *Curr Biol.* **19**: R198–R200.
 61. Scott, S.K., C. McGettigan & F. Eisner. 2009. A little more conversation, a little less action: candidate roles for motor cortex in speech perception. *Nature Rev. Neurosci.* **10**: 295–302.
 62. Fischer, J., C. Teufel, M. Drolet, *et al.* 2009. Orienting asymmetries and lateralized processing of sounds in humans. *BMC Neurosci.* **10**: 14.
 63. Wong, P.C., J.X. Jin, G.M. Gunasekera, *et al.* 2009. Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia* **47**: 693–703.
 64. Harris, K.C., J.R. Dubno, N.I. Keren, *et al.* 2009. Speech recognition in younger and older adults: a dependency on low-level auditory cortex. *J. Neurosci.* **29**: 6078–6087.
 65. Wiethoff, S., D. Wildgruber, W. Grodd & T. Ethofer. 2009. Response and habituation of the amygdala during processing of emotional prosody. *Neuroreport* **20**: 1356–1360.
 66. Fusar-Poli, P., A. Placentino, F. Carletti, *et al.* 2009. Laterality effect on emotional faces processing: ALE meta-analysis of evidence. *Neurosci. Lett.* **452**: 262–267.
 67. Herbert, C., T. Ethofer, S. Anders, *et al.* 2009. Amygdala activation during reading of emotional adjectives—an advantage for pleasant content. *Soc. Cogn. Affect. Neurosci.* **4**: 35–49.
 68. Wittfoth, M., C. Schröder, D.M. Schardt, *et al.* 2010. On emotional conflict: interference resolution of happy and angry prosody reveals valence-specific effects. *Cereb. Cortex*. **20**: 383–392.
 69. Kreifelts, B., T. Ethofer, T. Shiozawa, *et al.* 2009. Cerebral representation of non-verbal emotional perception: fMRI reveals audiovisual integration area between voice- and face-sensitive regions in the superior temporal sulcus. *Neuropsychologia*. **47**: 3059–3066.
 70. Ethofer, T., D. Van De Ville, K. Scherer & P. Vuilleumier. 2009. Decoding of emotional information in voice-sensitive cortices. *Curr. Biol.* **19**: 1028–1033.
 71. Obleser, J. & F. Eisner. 2009. Pre-lexical abstraction of speech in the auditory cortex. *Trends Cogn Sci.* **13**: 14–19.
 72. Hocking, J., K.L. McMahon & G.I. de Zubicaray. 2009. Semantic context and visual feature effects in object naming: an fMRI study using arterial spin labeling. *J. Cogn. Neurosci.* **21**: 1571–1583.
 73. Abel, S., K. Dressel, R. Bitzer, *et al.* 2009. The separation of processing stages in a lexical interference fMRI-paradigm. *Neuroimage* **44**: 1113–1124.
 74. Ali, N., D.W. Green, F. Kherif, *et al.* 2009. Role of the left caudate in suppressing irrelevant words. *J. Cogn. Neurosci.* doi:10.1162/jocn.2009.21352.
 75. Peck, K.K., J.F. Galgano, R.C. Branski, *et al.* 2009. Event-related functional MRI investigation of vocal pitch variation. *Neuroimage* **44**: 175–181.
 76. de Zubicaray, G. I. & K.L. McMahon. 2009. Auditory context effects in picture naming investigated with event-related fMRI. *Cogn. Affect. Behav. Neurosci.* **9**: 260–269.
 77. Tremblay, P. & V.L. Gracco. 2010. On the selection of words and oral motor responses: evidence of a response-independent fronto-parietal network. *Cortex*. **46**: 15–28.
 78. Hwang, K., E.D. Palmer, S. Basho, *et al.* 2009. Category-specific activations during word generation reflect experiential sensorimotor modalities. *Neuroimage* **48**: 717–725.
 79. Whitney, C., S. Weis, T. Krings, *et al.* 2009. Task-dependent modulations of prefrontal and hippocampal activity during intrinsic word production. *J. Cogn. Neurosci.* **21**: 697–712.
 80. Davis, M.H., A.M. Di Betta, M.J. Macdonald & M.G. Gaskell. 2009. Learning and consolidation of novel spoken words. *J. Cogn. Neurosci.* **21**: 803–820.
 81. Jeon, H.A., K.M. Lee, Y.B. Kim & Z.H. Cho. 2009. Neural substrates of semantic relationships: common and distinct left-frontal activities for generation of synonyms vs. antonyms. *Neuroimage* **48**: 449–457.

82. Heim, S., S.B. Eickhoff & K. Amunts. 2009. Different roles of cytoarchitectonic BA 44 and BA 45 in phonological and semantic verbal fluency as revealed by dynamic causal modelling. *Neuroimage* **48**: 616–624.
83. Fridriksson, J., D. Moser, J. Ryalls, *et al.* 2009. Modulation of frontal lobe speech areas associated with the production and perception of speech movements. *J. Speech Lang. Hear Res.* **52**: 812–819.
84. Zheng, Z.Z., K.G. Munhall & I.S. Johnsrude. 2009. Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *J. Cogn. Neurosci.* doi: 10.1162/jocn.2009.21324.
85. Heim, S., A.D. Friederici, N.O. Schiller, *et al.* 2009. The determiner congruency effect in language production investigated with functional MRI. *Hum. Brain Mapp.* **30**: 928–940.
86. Heim, S., S.B. Eickhoff, A.D. Friederici & K. Amunts. 2009. Left cytoarchitectonic area 44 supports selection in the mental lexicon during language production. *Brain Struct. Funct.* **213**: 441–456.
87. Brown, S., A.R. Laird, P. Q. Pfordresher, *et al.* 2009. The somatotopy of speech: phonation and articulation in the human motor cortex. *Brain Cogn.* **70**: 31–41.
88. Chang, S.E., M.K. Kenney, T.M. Loucks, *et al.* 2009. Common neural substrates support speech and non-speech vocal tract gestures. *Neuroimage* **47**: 314–325.
89. Kircher, T., K. Sass, O. Sachs & S. Krach. 2009. Priming words with pictures: neural correlates of semantic associations in a cross-modal priming task using fMRI. *Hum. Brain Mapp.* **30**: 4116–4128.
90. Melcher, T. & O. Gruber. 2009. Decomposing interference during Stroop performance into different conflict factors: an event-related fMRI study. *Cortex* **45**: 189–200.
91. Haupt, S., N. Axmacher, M. X. Cohen, *et al.* 2009. Activation of the caudal anterior cingulate cortex due to task-related interference in an auditory Stroop paradigm. *Hum. Brain Mapp.* **30**: 3043–3056.
92. Kotz, S.A., M. Schwartze & M. Schmidt-Kassow. 2009. Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. *Cortex* **45**: 982–990.
93. Shuster, L.I. 2009. The effect of sublexical and lexical frequency on speech production: an fMRI investigation. *Brain Lang.* **111**: 66–72.
94. Moser, D., J. Fridriksson, L. Bonilha, *et al.* 2009. Neural recruitment for the production of native and novel speech sounds. *Neuroimage* **46**: 549–557.
95. Byrd, K.E., L.M. Romito, M. Dzemidzic, *et al.* 2009. fMRI study of brain activity elicited by oral parafunctional movements. *J. Oral Rehabil.* **36**: 346–361.
96. Meister, I.G., D. Buelte, M. Staedtgen, *et al.* 2009. The dorsal premotor cortex orchestrates concurrent speech and fingertapping movements. *Eur. J. Neurosci.* **29**: 2074–2082.
97. Peschke, C., W. Ziegler, J. Kappes & A. Baumgaertner. 2009. Auditory-motor integration during fast repetition: the neuronal correlates of shadowing. *Neuroimage* **47**: 392–402.
98. van de Ven, V., F. Esposito & I.K. Christoffels. 2009. Neural network of speech monitoring overlaps with overt speech production and comprehension networks: a sequential spatial and temporal ICA study. *Neuroimage* **47**: 1982–1991.
99. Eickhoff, S.B., S. Heim, K. Zilles & K. Amunts. 2009. A systems perspective on the effective connectivity of overt speech production. *Philos. Transact. A Math. Phys. Eng. Sci.* **367**: 2399–2421.
100. Josse, G., F. Kherif, G. Flandin, *et al.* 2009. Predicting language lateralization from gray matter. *J. Neurosci.* **29**: 13516–13523.
101. Morgan, V.L., A. Mishra, A.T. Newton, *et al.* 2009. Integrating functional and diffusion magnetic resonance imaging for analysis of structure-function relationship in the human language network. *PLoS One* **4**: e6660.
102. Seghier, M.L. & C.J. Price. 2009. Reading aloud boosts connectivity through the putamen. *Cereb. Cortex.* doi: 10.1093/cercor/bhp123.
103. Levy, J., C. Pernet, S. Treserras, *et al.* 2009. Testing for the dual-route cascade reading model in the brain: an fMRI effective connectivity account of an efficient reading style. *PLoS One* **4**: e6675.
104. Londei, A., A. D'Ausilio, D. Basso, *et al.* 2009. Sensory-motor brain network connectivity for speech comprehension. *Hum. Brain Mapp.* doi: 10.1002/hbm.20888.
105. Heim, S., S.B. Eickhoff, A.K. Ischebeck, *et al.* 2009. Effective connectivity of the left BA 44, BA 45, and inferior temporal gyrus during lexical and phonological decisions identified with DCM. *Hum. Brain Mapp.* **30**: 392–402.
106. van Atteveldt, N., A. Roebroek & R. Goebel. 2009. Interaction of speech and script in human auditory cortex: Insights from neuro-imaging and effective connectivity. *Hear. Res.* **258**: 152–164.
107. Xiang, H.D., H.M. Fonteijn, D.G. Norris & P. Hagoort. 2009. Topographical functional connectivity pattern in the perisylvian language networks. *Cereb. Cortex.* doi: 10.1093/cercor/bhp119.

108. Liljeström, M., A. Hultén, L. Parkkonen & R. Salmelin. 2009. Comparing MEG and fMRI views to naming actions and objects. *Hum. Brain Mapp.* **30**: 1845–1856.
109. Vartiainen, J., T. Parviainen & R. Salmelin. 2009. Spatiotemporal convergence of semantic processing in reading and speech perception. *J. Neurosci.* **29**: 9271–9280.
110. Brem, S., P. Halder, K. Bucher, *et al.* 2009. Tuning of the visual word processing system: distinct developmental ERP and fMRI effects. *Hum. Brain Mapp.* **30**: 1833–1844.
111. Kotz, S.A. 2009. A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain Lang.* **109**: 68–74.
112. Pulvermüller, F., Y. Shtyrov & O. Hauk. 2009. Understanding in an instant: neurophysiological evidence for mechanistic language circuits in the brain. *Brain Lang.* **110**: 81–94.
113. D’Ausilio, A., F. Pulvermüller, P. Salmas, *et al.* 2009. The motor somatotopy of speech perception. *Curr. Biol.* **19**: 381–385.
114. Ruff, C.C., J. Driver & S. Bestmann. 2009. Combining TMS and fMRI: from ‘virtual lesions’ to functional-network accounts of cognition. *Cortex* **45**: 1043–1049.
115. Bestmann, S., C.C. Ruff, F. Blankenburg, *et al.* 2008. Mapping causal interregional influences with concurrent TMS-fMRI. *Exp. Brain Res.* **191**: 383–402.
116. Goebel, R. & N. van Atteveldt. 2009. Multisensory functional magnetic resonance imaging: a future perspective. *Exp. Brain Res.* **198**: 153–164.
117. Kiebel, S.J., K. von Kriegstein, J. Daunizeau & K.J. Friston. 2009. Recognizing sequences of sequences. *PLoS Comput. Biol.* **5**: e1000464.
118. Dominey, P.F. & T. Inui. 2009 Cortico-striatal function in sentence comprehension: insights from neurophysiology and modeling. *Cortex* **45**: 1012–1018.
119. Rajeev, D.S., R.D. Raizada, F.M. Tsao, *et al.* 2010. Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: prediction of individual differences. *Cereb. Cortex* . **20**: 1–12.
120. Hernandez, A.E. 2009. Language switching in the bilingual brain: what’s next? *Brain Lang.* **109**: 133–140.
121. Bloch, C., A. Kaiser, E. Kuenzli, *et al.* The age of second language acquisition determines the variability in activation elicited by narration in three languages in Broca’s and Wernicke’s area. *Neuropsychologia* **47**: 625–633.
122. Saur, D., A. Baumgaertner, A. Moehring, *et al.* 2009. Word order processing in the bilingual brain. *Neuropsychologia* **47** 158–168.
123. Chee, M.W. 2009. fMR-adaptation and the bilingual brain. *Brain Lang.* **109**: 75–79.
124. Abutalebi, J., M. Tettamanti & D. Perani. 2009. The bilingual brain: linguistic and non-linguistic skills. *Brain Lang.* **109**: 51–54.
125. Sakai, K.L., A. Nauchi, Y. Tatsuno, *et al.* 2009. Distinct roles of left inferior frontal regions that explain individual differences in second language acquisition. *Hum. Brain Mapp.* **30**: 2440–2452.
126. Weber, K. & P. Indefrey. 2009. Syntactic priming in German-English bilinguals during sentence comprehension. *Neuroimage* **46**: 1164–1172.
127. Zhang, Y., P.K. Kuhl, T. Imada, *et al.* 2009. Neural signatures of phonetic learning in adulthood: a magnetoencephalography study. *Neuroimage* **46**: 226–240.
128. Crinion, J.T., D.W. Green, R. Chung, *et al.* 2009. Neuroanatomical markers of speaking Chinese. *Hum. Brain Mapp.* **30**: 4108–4115.
129. Grogan, A., D.W. Green, N. Ali, *et al.* 2009. Structural correlates of semantic and phonemic fluency ability in first and second languages. *Cereb. Cortex.* **19**: 2690–2698.
130. Schnur, T.T., M.F. Schwartz, D.Y. Kimberg, *et al.* 2009. Localizing interference during naming: convergent neuroimaging and neuropsychological evidence for the function of Broca’s area. *Proc. Natl. Acad. Sci. USA* **106**: 322–327.
131. Kherif, F., G. Josse, M.L. Seghier & C.J. Price. 2009. The main sources of intersubject variability in neuronal activation for reading aloud. *J. Cogn.. Neurosci.* **21**: 654–668.
132. Seghier, M.L. & C.J. Price. 2009. Dissociating functional brain networks by decoding the between-subject variability. *Neuroimage* **45**: 349–259.
133. Richardson, F.M. & C.J. Price. 2009. Structural MRI studies of language function in the undamaged brain. *Brain Struct. Funct.* **213**: 511–523.
134. Flöel, A., M.H. de Vries, J. Scholz, *et al.* 2009. White matter integrity in the vicinity of Broca’s area predicts grammar learning success. *Neuroimage* **47**: 1974–1981.
135. Petkov, C.I., N.K. Logothetis & J. Obleser. 2009. Where are the human speech and voice regions, and do other animals have anything like them? *Neuroscientist* **15**: 419–429.