



Published in final edited form as:

*Phys Life Rev.* 2009 September 1; 6(3): 121–143. doi:10.1016/j.pprev.2009.06.001.

## The Functional Neuroanatomy of Language

**Gregory Hickok**

Center for Cognitive Neuroscience, Department of Cognitive Sciences, University of California, Irvine, Irvine, CA 92697, greg.hickok@uci.edu

### Abstract

There has been substantial progress over the last several years in understanding aspects of the functional neuroanatomy of language. These advances are summarized in this review. It will be argued that recognizing speech sounds is carried out in the superior temporal lobe bilaterally, that the superior temporal sulcus bilaterally is involved in phonological-level aspects of this process, that the frontal/motor system is not central to speech recognition although it may modulate auditory perception of speech, that conceptual access mechanisms are likely located in the lateral posterior temporal lobe (middle and inferior temporal gyri), that speech production involves sensory-related systems in the posterior superior temporal lobe in the left hemisphere, that the interface between perceptual and motor systems is supported by a sensory-motor circuit for vocal tract actions (not dedicated to speech) that is very similar to sensory-motor circuits found in primate parietal lobe, that verbal short-term memory can be understood as an emergent property of this sensory-motor circuit. These observations are understood within the context of a dual stream model of speech processing in which one pathway supports speech comprehension and the other supports sensory-motor integration. Additional topics of discussion include the functional organization of the planum temporale for spatial hearing and speech-related sensory-motor processes, the anatomical and functional basis of a form of acquired language disorder, conduction aphasia, the neural basis of vocabulary development, and sentence-level/grammatical processing.

### Language: what are we trying to understand?

Although human languages are often viewed colloquially as a family of communication codes developed thousands of years ago and passed down from one generation to the next, it is more accurate to think of human language as a computational system in the brain that computes a transform between thought (ideas, concepts, desires, etc.) on the one hand and an acoustic signal on the other<sup>1</sup>. For example, if one views an event, say a spotted dog biting a man, such an event can be understood, that is represented in some way in the conceptual system of the viewer. This conceptual representation, in turn, can be translated into any number of spoken utterances consisting of a stream of sounds such as *the dog bit the man*, *the pooch chomped his leg*, or *a Dalmatian attacked that gentleman*. The observation that there are many ways to express a concept through speech is a clue that the stuff of thought is not the same as spoken language. Pervasive ambiguity is another clue. *The stuffy nose* can be pronounced exactly the same as *the stuff he knows*; *we went to the bank at noon* could imply a financial transaction or a river-side picnic; and *I saw the man with binoculars* fails to indicate for sure who is in possession of the field glasses. Of course, this ambiguity exists in the speech signal, not in the

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

<sup>1</sup>Or in the case of signed languages, from thought to visuo-manual gestures.

mind of speaker, again telling us that language is not the same as thought. Rather, language is the means by which thoughts can be transduced into an acoustic code, and vice versa, albeit with an imperfect fidelity.

The transformation between thought and acoustic wave form and back again is a multistage computation. This is true even if we focus, as we will here, on processing stages that are likely computed “centrally,” that is, within the cerebral cortex.

On the output side, most researchers distinguish at *least* two major computational stages just in the production of a single word (Bock, 1999; Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Levelt, 1989). One stage involves selecting an appropriate *lexical item* (technically a *lemma*) to express the desired concept. Notice that there may be more than one lexical item to express a given concept (*dog, pooch, Dalmatian*), and in some cases there is no single lexical item to express a concept: there is no word for the top of a foot or the back of a hand; these concepts need to be expressed as phrases. At the lexical stage the system has access to information about the item’s meaning and grammatical category such as whether it is a noun or verb – information that is critical in structuring sentences -- but sound information is not yet available. One can think of this process as analogous to locating an entry in a reverse dictionary, one that is organized by meaning rather than alphabetically. The next stage involves accessing the sound structure, or *phonological form*, of the selected lexical item. Evidence for the cleavage between lexical and phonological processing comes from many sources, but speech production errors (slips of the tongue) are a rich source of data. Evidence for breakdowns at the lexical level come from word substitutions or exchanges such as George W. Bush’s utterance, *You’re working hard to put food on your family*, whereas breakdowns at the phonological level are evident in phonological exchanges such as the famous Spoonerism, *Is it kisstomary to cuss the bride?* (Dell, 1995; Fromkin, 1971). Beyond the single word level, production of connected speech requires additional computational stages involved in the construction of hierarchically organized sentences, intonational contours, and the like.

On the input side, again most researchers distinguish between at least two major stages just as in the processing of individual words, one involving the recovery of the phonological information (sound structure) and the other involving access to lexical-semantic information (Marslen-Wilson, 1987; McClelland & Elman, 1986).<sup>2</sup> It is an open question how much of these levels of processing may be shared between perception and production (Shelton & Caramazza, 1999). Some theorists posit additional processing levels in speech perception (Stevens, 2002) which may correspond to phonetic feature, segment, syllable, lexical, and semantic levels of processing. As with production, processing connected speech involves further computations related to parsing sentence structure and computing the compositional meaning coded in the relation between sentence structure and word meaning (e.g., the same words give rise to different utterance meanings depending on their position in a sentence structure: *the dog bit the man* vs. *the man bit the dog*).

In summary, the computational transformation between thought and acoustic waveform (and vice versa) is a complicated, multistage process. Not only does it involve several stages of processing in the sensory and motor periphery (not discussed here), but also involves multiple linguistic stages within central brain systems. The object of much research on the neuroscience of language involves mapping the neural circuits that support the various levels and stages of these computational transformations, and understanding the relation between input and output systems, as well as to related non-linguistic functions.

---

<sup>2</sup>An entirely different approach, the motor theory of speech perception (Lieberman & Mattingly, 1985) will be discussed below.

The balance of this review will summarize what is currently understood regarding the neural basis of speech sound recognition, the neural territory involved in accessing meaning from acoustically presented speech, the role of sensory systems in speech production, the neural circuits underlying sensory-motor integration for speech and related functions, verbal short-term memory, and the organization of the planum temporale, a region classically associated with speech functions. We will then discuss a dual stream framework within which these findings can be understood. Finally, we will discuss the neural basis of higher-order aspects of language processing.

## Speech recognition is bilaterally organized in the superior temporal gyrus

The transformation from the acoustic speech signal into a conceptual representation involves several processing steps in the auditory periphery that will not be discussed here. We pick up the story at the level of auditory cortex and consider the systems involved in processing speech sounds during spoken word recognition. Figure 1 shows some of the relevant anatomy and functional organization that will be discussed throughout the review.

In contrast to common assumptions regarding the functional anatomy of language, evidence from a variety of sources indicates that phonological stages of spoken word recognition are supported by neural systems in the superior temporal lobe -- superior temporal gyrus (STG) and superior temporal sulcus (STS) --*bilaterally*. In neuroimaging studies, listening to speech activates the superior temporal lobe bilaterally and largely symmetrically (Binder et al., 2000; Binder et al., 1994; Mazoyer et al., 1993; Price et al., 1996; Schlosser, Aoyagi, Fulbright, Gore, & McCarthy, 1998; Zatorre, Meyer, Gjedde, & Evans, 1996). Such a finding, however, does not tell us which aspect of speech recognition may be processed in the two hemispheres: it is possible that while activation in spoken word recognition is bilateral, phonological stages of processing are nonetheless restricted to the left hemisphere. This hypothesis predicts that damage to the left superior temporal lobe should produce profound phonological deficits in spoken word recognition. However, this is not the case. Damage to the posterior superior temporal lobe, such as in certain forms of aphasia -- language disorders caused by brain injury, typically stroke -- (A. R. Damasio, 1991; Damasio, 1992) *does* produce deficits in spoken word recognition (Goodglass, 1993; Goodglass, Kaplan, & Barresi, 2001), however, these deficits involve only mild phonological processing impairment, and in fact, appear to result predominantly from a disruption to lexical-semantic level processes (Bachman & Albert, 1988; Baker, Blumstein, & Goodglass, 1981; Gainotti, Micelli, Silveri, & Villa, 1982; Miceli, Gainotti, Caltagirone, & Masullo, 1980). This conclusion is based on experiments in which patients are presented with a spoken word and asked to point to a matching picture within an array that includes phonological, semantic, and unrelated foils; phonological error rates are low overall (5–12%) with semantic errors dominating. This tendency also holds in acute aphasia (Breese & Hillis, 2004; Rogalsky, Pitz, Hillis, & Hickok, 2008) showing that the relative preservation of phonological abilities in unilateral aphasia is not a result of long-term plastic reorganization. Data from split-brain patients (individuals who have undergone surgical cutting of their corpus callosum to treat epilepsy) (Zaidel, 1985) and Wada procedures (Hickok et al., 2008; McGlone, 1984) (a pre-neurosurgical procedure in which one and then the other cerebral hemisphere is anesthetized to assess the lateralization of language and memory functions (Wada & Rasmussen, 1960)) also indicate that the right hemisphere alone is capable of good auditory comprehension at the word level, and that when errors occur, they are more often semantic than phonological (Figure 2).

Disruption of the left superior temporal lobe does not lead to severe impairments in phonological processing during spoken word recognition. This observation has led to the hypothesis that phonological processes in speech recognition are bilaterally organized in the superior temporal lobe (Hickok & Poeppel, 2000, 2004, 2007). Consistent with this claim is

the observation that damage to the STG bilaterally produces profound impairment in spoken word recognition, in the form of word deafness, a condition in which basic hearing is preserved (pure tone thresholds within normal limits) but the ability to comprehend speech is effectively nil (Buchman, Garron, Trost-Cardamone, Wichter, & Schwartz, 1986).

## The superior temporal sulcus is a critical site for phonological processing

The STS has emerged as an important site for representing and/or processing phonological information (Binder et al., 2000; Hickok & Poeppel, 2004, 2007; Indefrey & Levelt, 2004; Liebenthal, Binder, Spitzer, Possing, & Medler, 2005; Price et al., 1996). Functional imaging studies designed to isolate phonological processes in perception by contrasting speech stimuli with complex non-speech signals have found activation along the STS (Liebenthal et al., 2005; Narain et al., 2003; Obleser, Zimmermann, Van Meter, & Rauschecker, 2006; Scott, Blank, Rosen, & Wise, 2000; Spitsyna, Warren, Scott, Turkheimer, & Wise, 2006; Vouloumanos, Kiehl, Werker, & Liddle, 2001). Other investigators have manipulated psycholinguistic variables that tap phonological processing systems as a way of identifying phonological networks. This approach also implicates the STS (Okada & Hickok, 2006). Although many authors consider this system to be strongly left dominant, both lesion and imaging evidence suggest a bilateral organization (Figure 3) (Hickok & Poeppel, 2007).

One currently unresolved question is the relative contribution of anterior versus posterior STS regions in phonological processing. Lesion evidence indicates that damage to posterior temporal lobe areas are most predictive of auditory comprehension deficits (Bates et al., 2003), however, as noted above, comprehension deficits in aphasia result predominantly from post-phonemic processing levels. A majority of functional imaging studies targeting phonological processing in perception have highlighted regions in the posterior half of the STS (Figure 3) (Hickok & Poeppel, 2007). Other studies, however, have reported *anterior* STS activation in perceptual speech tasks (Mazoyer et al., 1993; Narain et al., 2003; Scott et al., 2000; Spitsyna et al., 2006). These studies involved sentence-level stimuli raising the possibility that anterior STS regions may be responding to some other aspect of the stimuli such as its syntactic or prosodic organization (Friederici, Meyer, & von Cramon, 2000; Humphries, Binder, Medler, & Liebenthal, 2006; Humphries, Love, Swinney, & Hickok, 2005; Humphries, Willard, Buchsbaum, & Hickok, 2001; Vandenberghe, Nobre, & Price, 2002). The weight of the available evidence, therefore, suggests that the critical portion of the STS that is involved in phonological-level processes, is approximately bounded anteriorly by the anterolateral-most aspect of Heschl's gyrus and posteriorly by the posterior-most extent of the Sylvian fissure (see Figure 1 & 6) (Hickok & Poeppel, 2007).

## Role of motor systems in speech recognition

The evidence outlined above points to auditory-related areas in the temporal lobe as the critical substrate for speech recognition. An alternative view that is getting a lot of attention recently is that the motor speech system plays a critical role in speech recognition (D'Ausilio et al., 2009; Fadiga & Craighero, 2006; Fadiga, Craighero, Buccino, & Rizzolatti, 2002). This is not a new idea and in fact dates back to the 1960s when the motor theory of speech perception was proposed by Liberman and colleagues (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). The motivation for this proposal was the observation that there is not a one-to-one correspondence between acoustic patterns and perceived phonemes. For example, the acoustics of a "d" sound is different depending on whether that sound appears in the context of the syllable /di/ versus the syllable /da/ (Figure 4). Liberman et al. noticed that despite the acoustic differences, the manner in which /d/ is *articulated* is invariant (this claim has been subsequently questioned). Therefore, they proposed that "the objects of speech perception are the intended

phonetic gestures of the speaker, represented in the brain as invariant motor commands” (Liberman & Mattingly, 1985) (p. 2).

Empirical work aimed at testing this hypothesis failed to support the motor theory of speech perception such that by the 1990s the claim had few supporters among speech scientists, a situation that remains true today (Galantucci, Fowler, & Turvey, 2006). However, the discovery of mirror neurons in the frontal lobe of macaque monkeys revitalized the theory, at least among neuroscientists if not speech scientists (Galantucci et al., 2006; Lotto, Hickok, & Holt, 2009). Mirror neurons respond both during action execution and action perception (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996), a response pattern that had led to the hypothesis that “... we understand action because the motor representation of that action is activated in our brain” (Rizzolatti, Fogassi, & Gallese, 2001)(p. 661). In these models, speech perception is often included as a form of action perception (Fadiga & Craighero, 2006; Gallese et al., 1996; Rizzolatti & Arbib, 1998).

There is good evidence that motor-related systems and processes can play a role in at least some speech perception tasks (Galantucci et al., 2006). For example, stimulating motor lip areas produces a small bias to hear partially ambiguous speech sounds as a sound formed by lip closure (e.g., /b/) whereas stimulation of motor tongue areas bias perception toward sounds with prominent tongue movements (e.g., /t/) (D’Ausilio et al., 2009). But such findings do not necessarily mean that motor representations are central to speech sound recognition. Instead, motor-related processes may simply modulate auditory speech recognition systems in a top-down fashion (Hickok, 2009). The stronger claim, that motor systems are central to the process of speech recognition can be assessed by examining the neuropsychology literature: if motor speech systems are critical for speech recognition, then one should find evidence that damage to motor speech systems produces deficits in speech recognition.

Several lines of evidence argue against this view. First, while focal damage restricted to the region of primary motor cortex that projects to face and mouth muscles can produce acute severe speech production deficits (e.g., output limited to grunts), it can leave speech comprehension “completely normal” (Terao et al., 2007)(p. 443). Second, severe chronic “Broca’s aphasia” caused by large lesions involving Broca’s region, lower primary motor cortex and surrounding tissue can leave the patient with no speech output or with only stereotyped output, but with relatively preserved comprehension (Naeser, Palumbo, Helm-Estabrooks, Stiassny-Eder, & Albert, 1989). Third, even acute deactivation of the entire left hemisphere in patients undergoing Wada procedures, which produces complete speech arrest, leaves speech sound perception relatively intact (phonemic error rate < 10%) (Hickok et al., 2008). This pattern holds even when fine phonetic discrimination is required for successful comprehension (e.g., comprehending *bear* vs. *pear* which only differ by one feature) (Hickok et al., 2008). Fourth, bilateral lesions to the cortex of the anterior operculum (region buried in the Sylvian fissure beneath blue shaded region in Figure 1) can cause anarthria, that is, loss of voluntary muscle control of speech, yet these lesions do not cause speech recognition deficits (Weller, 1993). Fifth, *bilateral* lesions to Broca’s area, argued to be the core of the human mirror system (Craighero, Metta, Sandini, & Fadiga, 2007), do not cause word level speech recognition deficits (Levine & Mohr, 1979). Sixth, the failure of a child to develop motor speech ability, either as a result of a congenital anarthria (Lenneberg, 1962) or an acquired anarthria secondary to bilateral anterior operculum lesions (Christen et al., 2000) do not preclude the development of normal receptive speech. Seventh, babies develop sophisticated speech perception abilities including the capacity to make fine distinctions and perceive speech categorically as early as 1-month of age, well before they develop the ability to produce speech (Eimas, Siqueland, Jusczyk, & Vigorito, 1971). Finally, species without the capacity to develop speech (e.g., chinchillas) can nonetheless be trained to perceive subtle speech-sound

distinctions in a manner characteristic of human listeners, that is, categorically (Kuhl & Miller, 1975).

In short, disruption of number of levels of the motor-speech system, including its complete failure to develop, does not preclude the ability to make subtle speech sound discriminations in perception. This suggests that to the extent that motor systems influence the perception of speech, they do so in a modulatory fashion and not as a central component in speech recognition systems. Put differently, speech perception is an auditory phenomenon, but one that can be modulated by other sources of information including, but not limited to, the motor speech system.

## Phonological processing systems in speech recognition are bilateral but asymmetric

The hypothesis that phoneme-level processes in speech recognition are bilaterally organized does not imply that the two hemispheres are computationally identical. In fact there is strong evidence for hemispheric differences in the processing of acoustic/speech information (Abrams, Nicol, Zecker, & Kraus, 2008; Boemio, Fromm, Braun, & Poeppel, 2005; Giraud et al., 2007; Hickok & Poeppel, 2007; Zatorre, Belin, & Penhune, 2002). What is the basis of these differences? One view is that the difference turns on selectivity for temporal (left hemisphere) versus spectral (right hemisphere) resolution (Zatorre et al., 2002). Another proposal is that the two hemispheres differ in terms of their sampling rate, with the left hemisphere operating at a faster rate (25–50 Hz) and the right hemisphere at a slower rate (4–8 Hz) (Poeppel, 2003)<sup>3</sup>. Further research is needed to sort out these details. For present purposes, however, what is important to note is that this asymmetry of function indicates that spoken word recognition involves parallel pathways (multiple routes) in the mapping from sound to meaning (Figure 5) (Hickok & Poeppel, 2007). Although this conclusion differs from standard models of speech recognition (Luce & Pisoni, 1998; Marslen-Wilson, 1987; McClelland & Elman, 1986), it is consistent with the fact that speech contains redundant cues to phonemic information, and with behavioral evidence suggesting that the speech system can take advantage of these different cues (Remez, Rubin, Pisoni, & Carrell, 1981; Shannon, Zeng, Kamath, Wygonski, & Ekelid, 1995).

## Accessing conceptual semantic systems

The neural organization of conceptual-semantic systems is a matter of debate. A common view is that conceptual information is represented in a widely distributed fashion throughout cortex and that these representations involve the same sensory, motor, and supramodal cortical systems originally invoked in processing that information (Damasio, 1989; Martin, 1998; Martin & Chao, 2001; Pulvermuller, 1996; Wernicke, 1874/1969). Other researchers argue for a more focally organized semantic “hub” in the anterior temporal region (Patterson, Nestor, & Rogers, 2007), and still others hold that semantic knowledge is organized into functionally specialized neural systems dedicated to processing information from evolutionarily relevant conceptual categories, i.e., domains of knowledge that have survival and therefore reproductive value such as animals, fruits/vegetables, and possibly tools (Caramazza & Mahon, 2003). These are interesting and complex issues. The question we will address in this section is more restricted, however. Assuming there are systems in the superior temporal lobe (bilaterally) that process acoustic and phonemic-level information during speech recognition, and assuming that conceptual-semantic information involves cortical regions outside of the superior temporal lobe (an uncontroversial assertion), then how do these two types of information come together?

---

<sup>3</sup>These two proposals are not incompatible as there is a relation between sampling rate and spectral vs. temporal resolution (Zatorre et al., 2002).

Clearly some association must be established, but the neural mechanism remains to be specified. At most, existing evidence points to the posterior lateral and inferior temporal regions (middle and inferior temporal gyri, posterior pink shaded region in Figure 1) as important in mapping sound onto meaning (Hickok & Poeppel, 2004, 2007).

Damage to posterior temporal lobe regions, particularly along the middle temporal gyrus, has long been associated with auditory comprehension deficits (Bates et al., 2003; H. Damasio, 1991; Dronkers, Redfern, & Knight, 2000). We can infer that these deficits are predominantly post-phonemic in nature because auditory comprehension deficits secondary to unilateral lesions tend to produce relatively mild phonemic perception deficits (see above). A handful of stroke-induced cases of relatively pure “semantic” level deficits in word processing have been linked to the posterior lateral and inferior temporal lobe (Chertkow, Bub, Deaudon, & Whitehead, 1997; Hart & Gordon, 1990). Data from direct cortical stimulation studies corroborate the involvement of the middle temporal gyrus in auditory comprehension, but also indicate the involvement of a much broader network involving most of the superior temporal lobe (including anterior portions), and the inferior frontal lobe (Miglioretti & Boatman, 2003). Functional imaging studies have also implicated posterior middle temporal regions in lexical-semantic processing (Binder et al., 1997; Rissman, Eliassen, & Blumstein, 2003; Rodd, Davis, & Johnsrude, 2005).

Anterior temporal lobe (ATL) regions (anterior pink shaded region in Figure 1) have also been implicated both in lexical-semantic and sentence-level processing (syntactic and/or semantic integration processes). Patients with semantic dementia, a condition that has been used to argue for a lexical-semantic function of the ATL (Scott et al., 2000; Spitsyna et al., 2006), have atrophy involving the ATL bilaterally, along with deficits on lexical tasks such as naming, semantic association, and single-word comprehension (Gorno-Tempini et al., 2004; Hodges & Patterson, 2007). However, these deficits might be more general, given that the atrophy involves a number of regions in addition to the lateral ATL, including bilateral inferior and medial temporal lobe (known to support aspects of memory function), bilateral caudate nucleus, and right posterior thalamus, among others (Gorno-Tempini et al., 2004). Furthermore, these deficits appear to be amodal, affecting the representation of semantic knowledge for objects generally (Patterson et al., 2007), rather than a process limited to mapping between sound and meaning. Specifically, patients with semantic dementia have difficulty accessing object knowledge not only from auditory language input, but also from visual input showing that the deficit is not restricted to sensory-conceptual mapping within one modality, but instead affects supramodal representations of conceptual knowledge of objects (Patterson et al., 2007). One hypothesis then is that the posterior lateral and inferior temporal lobe (posterior pink region in Figure 1) seems to be more involved in accessing semantic knowledge from acoustic input, whereas the anterior temporal lobe seems to be more involved in integrating certain forms of semantic knowledge *across* modalities (Patterson et al., 2007).

## Posterior language cortex in the left hemisphere is involved in phonological aspects of speech production

While there is little evidence for a strong motor component in speech *perception* (see above), there is unequivocal evidence for an important influence of sensory systems on speech *production*. Behaviorally, it is well-established that auditory input can produce rapid and automatic effects on speech production. For example, delayed auditory feedback of one’s own voice disrupts speech fluency (Stuart, Kalinowski, Rastatter, & Lynch, 2002; Yates, 1963). Other forms of altered speech feedback have similar effects: shifting the pitch or first formant (frequency band of speech) in the auditory feedback of a speaker results in rapid compensatory modulation of speech output (Burnett, Senner, & Larson, 1997; Houde & Jordan, 1998).

Finally, adult onset deafness is associated with articulatory decline (Waldstein, 1989), indicating that auditory feedback is important in maintaining articulatory tuning.

Sensory guidance of speech production is not limited to phonetic or pitch level processes, but also supports phonemic sequence production. Consider the fact that listeners can readily reproduce a novel sequence of phonemes. If presented auditorily with a non-word, say *nederop*, one has no difficulty reproducing this stimulus vocally. In order to perform such a task, the stimulus must be coded in the auditory/phonological system and then mapped onto a corresponding motor/articulatory sequence. An auditory-motor integration system must therefore exist to perform such a mapping at the level of sound *sequences* (Doupe & Kuhl, 1999; Hickok, Buchsbaum, Humphries, & Muftuler, 2003; Hickok & Poeppel, 2000, 2004, 2007). This system is critical for new vocabulary development (including second language acquisition) where one must learn the sound pattern of new words along with their articulatory sequence patterns (see below for additional discussion), but may also be involved in sensory guidance of the production of infrequent and/or high phonological load word forms. For example, compare the articulatory agility in pronouncing technical terms -- such as *lateral geniculate nucleus* or *arcuate fasciculus* -- between students who have learned such terms but do not have much experience articulating them, and experts/lecturers on gross neuroanatomy who have much experience articulating these terms. In the former case, articulation is slow and deliberate as if it were being guided syllable by syllable, whereas in the latter it is rapid and automatic. One explanation for this pattern is that for the novice, a complex sequence must be guided by the learned sensory-representation of the word form, whereas for the expert the sequence has been “chunked” as a motor unit requiring little sensory guidance, much as a practiced sequence of keystrokes can be chunked and automated.

Given these behavioral observations, it is no surprise that posterior sensory-related cortex in the left hemisphere have been found to play an important role in speech production. For example, damage to the left dorsal posterior superior temporal gyrus and/or the supramarginal gyrus (Figure 1) is associated with speech production deficits. In particular, such a lesion is associated with conduction aphasia, a syndrome typically caused by stroke that is characterized by good auditory comprehension, but frequent phonemic errors in speech production, naming difficulties that often involve tip-of-the-tongue states (implicating a breakdown in phonological encoding), and difficulty with verbatim repetition (Baldo, Klostermann, & Dronkers, 2008; Damasio & Damasio, 1980; Goodglass, 1992)<sup>4</sup>.

Conduction aphasia has classically been considered to be a disconnection syndrome involving damage to the arcuate fasciculus, the white matter pathway that connects to the posterior superior temporal lobe with the posterior inferior frontal lobe (Geschwind, 1965). However, there is now good evidence that conduction aphasia results from cortical dysfunction (Anderson et al., 1999; Baldo et al., 2008; Hickok et al., 2000). The production deficit in conduction aphasia is load-sensitive: errors are more likely on longer, lower-frequency words, and verbatim repetition of strings of speech with little semantic constraint (Goodglass, 1992; Goodglass, 1993). Thus, conduction aphasia provides evidence for the involvement of left posterior auditory-related brain regions in phonological aspects of speech production. See also (Wise et al., 2001).

Functional imaging evidence also implicates left posterior superior temporal regions in speech production generally (Hickok et al., 2000; Price et al., 1996), and phonological stages of the

---

<sup>4</sup>Although conduction aphasia is often characterized as a disorder of repetition, it is clear that the deficit extends well beyond this one task (Hickok et al., 2000). In fact, Wernicke first identified conduction aphasia as a disorder of speech production in the face of preserved comprehension (Wernicke, 1874/1969). It was only later that Lichtheim introduced repetition as a convenient diagnostic tool for assessing the integrity of the link between sensory and motor speech systems (Lichtheim, 1885).

process in particular (Indefrey & Levelt, 2004; Indefrey & Levelt, 2000). With respect to the latter, the posterior portion of the left planum temporale region (Figure 6), which is within the distribution of lesions associated with conduction aphasia, activates during picture naming and exhibits length effects (Okada, Smith, Humphries, & Hickok, 2003), frequency effects (Graves, Grabowski, Mahta, & Gordon, 2007), and has a time-course of activation, measured electromagnetically, that is consistent with the phonological encoding stage of naming (Levelt, Praamstra, Meyer, Helenius, & Salmelin, 1998).

Taken together, the lesion and physiological evidence reviewed in this section make a compelling argument for the involvement of left posterior superior temporal regions in phonological aspects of speech production.

## **A sensory-motor integration network for the vocal tract**

What is the nature of the sensory-motor integration circuit for speech? In classical models a sensory-motor speech connection was instantiated as a simple white matter pathway, the arcuate fasciculus (Geschwind, 1971). More recent proposals have argued, instead, for a *cortical* system that serves to integrate sensory and motor aspects of speech, that is, they perform a transform between auditory representations of speech and motor representations of speech (Hickok et al., 2003; Hickok et al., 2000; Hickok & Poeppel, 2000, 2004, 2007; Warren, Wise, & Warren, 2005). The proposal is analogous to sensory-motor integration regions found in the dorsal stream of the primate visual system (Andersen, 1997; Colby & Goldberg, 1999; Milner & Goodale, 1995).

Some background on sensory-motor integration areas in the macaque monkey will be instructive here. In the parietal lobe of the macaque, the intraparietal sulcus (IPS) contains a constellation of functional regions that support sensory-motor integration (Andersen, 1997; Colby & Goldberg, 1999; Grefkes & Fink, 2005) (Figure 7). These regions are organized around motor effector systems. For example, area AIP supports sensory-motor integration for grasping – cells respond both to viewing graspable objects and to grasping actions, and the region is connected to frontal area F5 which is involved in grasping -- whereas area LIP supports sensory-motor integration for eye movements – cells respond to detection of a visual target as well as during eye movements toward that target, and the region is connected to the frontal eye field (for a recent review see (Grefkes & Fink, 2005)). These areas receive multi-sensory input indicating that they are not specifically tied to a particular sensory processing stream but rather to any source of sensory information that may be useful in guiding actions of a given motor effector system. Human homologues of these monkey IPS regions have been delineated (Grefkes & Fink, 2005).

A series of studies over the last several years has identified a cortical network for speech and related abilities (e.g., vocal music) which has many of the properties exhibited by sensory-motor networks in the macaque IPS including sensory-motor response properties, connectivity with frontal motor systems, motor-effector specificity, and multisensory responses. The speech-related network with these response properties includes an area (termed Spt, Sylvian parietal-temporal) in the left posterior planum temporale region (Figure 1), that has been argued to support sensory-motor integration for speech (Hickok et al., 2003). Because of similarities between the response properties of area Spt and IPS areas, Spt has been proposed as a sensory-motor integration area for the vocal tract effector (Hickok, Okada, & Serences, 2009; Pa & Hickok, 2008). We will review the evidence for this claim below.

### **Spt exhibits sensory-motor response properties**

A number of fMRI studies have demonstrated the existence of an area in the left posterior Sylvian region (area Spt, Figure 8A) that responds both during the perception and production

of speech (Figure 8B), even when speech is produced covertly (subvocally) so that there is no overt auditory feedback (Buchsbaum, Hickok, & Humphries, 2001; Buchsbaum, Olsen, Koch, & Berman, 2005; Buchsbaum, Olsen, Koch, Kohn et al., 2005; Hickok et al., 2003). Spt is not speech-specific, however. It responds equally well to the perception and (covert) production via humming of melodic stimuli (Figure 8B) (Hickok et al., 2003; Pa & Hickok, 2008). A recent study has demonstrated that the pattern of activity across voxels in Spt is different during the sensory and motor phase of such tasks indicating partially distinct populations of cells, some sensory-weighted and some motor-weighted (Figure 8C)(Hickok et al., 2009). A similar distribution of cell types has been found in monkey IPS sensory-motor areas (Grefkes & Fink, 2005).

### **Spt is functionally connected to motor speech areas**

Spt activity is tightly correlated with activity in frontal speech-production related areas, such as the pars opercularis (BA 44) (Buchsbaum et al., 2001) suggesting that the two regions are functionally connected. Furthermore, cortex in the posterior portion of the planum temporale (area Tpt; Figure 5)) has a cytoarchitectonic structure (cellular layering pattern in cortex) that is similar to BA44. Galaburda writes, area Tpt "...exhibits a degree of specialization like that of Area 44 in Broca's region. It contains prominent pyramids in layer IIIc and a broad lamina IV.... the intimate relationship and similar evolutionary status of Areas 44 and Tpt allows for a certain functional overlap" (Galaburda, 1982).

### **Spt activity is modulated by motor effector manipulations**

In monkey parietal cortex sensory-motor integration areas are organized around motor effector systems. Recent evidence suggests that Spt is organized around the vocal tract effector system: Spt was less active when skilled pianists listened to and then imagined playing novel melodies than when they listened to and covertly hummed the same melodies (Figure 8D)(Pa & Hickok, 2008).

### **Spt is sensitive to speech-related visual stimuli**

Many neurons in sensory-motor integration areas of the monkey parietal cortex are sensitive to inputs from more than one sensory modality (Andersen, 1997; Grefkes & Fink, 2005). The planum temporale, while often thought to be an auditory area, also activates in response to sensory input from other modalities. For example, silent lip-reading has been shown to activate auditory cortex in the vicinity of the planum temporale (Calvert et al., 1997; Calvert & Campbell, 2003). Although these studies typically report the location as "auditory cortex" including primary regions, group-based localizations in this region can be unreliable. Indeed a recent fMRI study found that activation to visual speech and activation using the standard Spt-defining auditory-motor task (listen then covertly produce) overlap (Okada & Hickok, 2009). Further, human cytoarchitectonic studies (Galaburda & Sanides, 1980) and comparative studies in monkeys (Smiley et al., 2007) indicate that the posterior PT region is not part of unimodal auditory cortex. Thus, Spt appears to be multisensory rather than part of auditory cortex proper.

### **Damage to Spt produces sensory-motor deficits but not speech recognition deficits**

While Spt is responsive to speech stimulation, it is not critical for speech recognition just as IPS sensory-motor areas are not critical for object recognition (Milner & Goodale, 1995; Ungerleider & Mishkin, 1982). Damage to the posterior Sylvian parietal-temporal region is associated with conduction aphasia (Baldo et al., 2008), a syndrome characterized by phonological errors in speech production and difficulty with verbatim repetition of heard speech. Speech recognition is well preserved at the word level in conduction aphasia. For this

reason, conduction aphasia has been argued to be a deficit in sensory-motor integration for speech (Hickok, 2000; Hickok et al., 2003; Hickok & Poeppel, 2004).

Thus Spt exhibits all the features of sensory-motor integration areas as identified in the parietal cortex of the monkey. This suggests that Spt is a sensory-motor integration area for vocal tract actions (Hickok et al., 2009; Pa & Hickok, 2008), placing it in the context of a network of sensory-motor integration areas in the posterior parietal and temporal/parietal cortex, which receive multisensory input and are organized around motor-effector systems. Although area Spt is not language-specific, it counts sensory-motor integration for phonological information as a prominent function.

Additional sources of data provide further evidence for the role of Spt in sensory guidance of speech production. It was noted above that altered auditory feedback provides strong evidence for the role of the sensory systems in speech production. A recent functional imaging study has shown that Spt activity is increased during altered feedback relative to non-altered feedback (Tourville, Reilly, & Guenther, 2008) suggesting a role for Spt in this form of auditory-motor integration. Additionally, conduction aphasics, who typically have damage involving Spt, have been reported to exhibit a decreased sensitivity to the disruptive effects of delayed auditory feedback (Boller & Marcie, 1978; Boller, Vrtunski, Kim, & Mack, 1978). Further circumstantial evidence comes from stutterers. Stutterers sometimes exhibit a paradoxical improvement in fluency under conditions of altered auditory feedback (Stuart, Frazier, Kalinowski, & Vos, 2008) implicating some form of auditory-motor integration anomaly in the etiology of the disorder. A recent study has found that this paradoxical response to delayed auditory feedback is correlated with an atypical asymmetry of the planum temporal (Foundas et al., 2004); Spt is found within this region.

In summary, several diverse lines of evidence point to Spt as a region supporting sensory-motor integration for vocal tract actions.

## Relation between verbal short-term memory and sensory-motor integration

Verbal short-term memory is typically characterized by a storage component and a mechanism for active maintenance of this information. In Baddeley's "phonological loop" model the storage mechanism is the "phonological store," a dedicated buffer, and active maintenance is achieved by the "articulatory rehearsal" mechanism (Figure 9A) (Baddeley, 1992). The concept of a sensory-motor integration network as outlined above, provides an independently motivated neural circuit that may be the basis for verbal short-term memory (Buchsbaum, Olsen, Koch, & Herman, 2005; Hickok et al., 2003; Hickok & Poeppel, 2000) (see also (Aboitiz & García V., 1997; Jacquemot & Scott, 2006)). Specifically, on the assumption that the proposed sensory-motor integration circuit is bidirectional (Hickok & Poeppel, 2000, 2004, 2007), one can equate the storage component of verbal short-term memory with phonological processing systems in the superior temporal lobe (the same STS regions that are involved in sensory/recognition processes), and one can equate the active maintenance component with frontal articulatory systems: the sensory-motor integration network (Spt) allows articulatory mechanisms to maintain verbal information in an active state (Figure 9B) (Hickok et al., 2003). In this sense, the basic architecture is similar to Baddeley's, except that there is a proposed computational mechanism (sensory-motor transformations in Spt) mediating the relation between the storage and active maintenance components. This view differs from Baddeley's however, in that it assumes that the storage component is not a dedicated buffer, but an active state of networks that are involved in perceptual recognition (Fuster, 1995; Ruchkin, Grafman, Cameron, & Berndt, 2003). Because evidence suggests that the sensory-motor integration network is not specific to phonological information (Hickok et al., 2003), we also suggest that the verbal short-term memory circuit is not specific to phonological

information, a position that is in line with recent behavioral work (Jones, Hughes, & Macken, 2007; Jones & Macken, 1996; Jones, Macken, & Nicholls, 2004).

Conduction aphasia has been characterized as a disorder of verbal working memory in which the functional damage involves the phonological buffer (Baldo et al., 2008). Evidence in favor of this claim comes from the observation that conduction aphasics comprehend the gist of an utterance, but fail to retain the phonological details. For example, patients may paraphrase a sentence they attempt to repeat (Baldo et al., 2008). However, this view fails to account for other typical symptoms of conduction aphasia such as phonological errors in speech production. A sensory-motor integration account can capture the broader symptom complex because such a mechanism participates in sensory guidance of speech production as well as supporting verbal short-term memory (Hickok et al., 2003; Hickok & Poeppel, 2004). For a thorough discussion of the relation between verbal working memory and conduction aphasia see (Buchsbaum & D'Esposito, 2008).

I suggested above that the proposed sensory-motor integration network supports the acquisition of new vocabulary. It is relevant in this context that Baddeley's phonological loop model of verbal short-term memory has been correlated with word learning in children and adults (Baddeley, Gathercole, & Papagno, 1998). To the extent that the phonological loop can be explained in terms of a sensory-motor integration network, the association between word learning and phonological short-term memory provides empirical support for the role of sensory-motor circuits in vocabulary acquisition. In support of this possibility recent studies of novel word learning have implicated area Spt as an important region in the acquisition of new phonological word forms (McNealy, Mazziotta, & Dapretto, 2006; Paulesu et al., 2009).

## **Functional anatomic housekeeping: the organization of the planum temporale**

Area Spt, a critical node in the proposed sensory-motor network for the vocal tract articulators is located in the planum temporale (PT) region in the left hemisphere (Figure 6). This region has long been associated with speech functions as a result of (i) the discovery that the left PT is larger than the right PT in most individuals (Geschwind & Levitsky, 1968), and (ii) the proximity of the PT to classical Wernicke's area. However, research in the 1990s challenged the view that the PT is strictly a speech area. Functional imaging work showed that the left PT responds better to tones than to speech, at least in some tasks (Binder, Frost, Hammeke, Rao, & Cox, 1996). The leftward PT asymmetry was found to be correlated with musical ability (Schlaug, Jancke, Huang, & Steinmetz, 1995). It was discovered that chimpanzee's have the same leftward asymmetry of the PT (Gannon, Holloway, Broadfield, & Braun, 1998) showing that the pattern is not speech-driven. And the PT was found to be activated by a range of non-speech stimuli including aspects of spatial hearing (Smith, Okada, Saberi, & Hickok, 2004; Warren & Griffiths, 2003; Warren, Zielinski, Green, Rauschecker, & Griffiths, 2002). This multi-functionality of the PT led some researchers to propose that the PT acts a "computational hub" in which acoustic signals of a variety of sorts serve as input and via a pattern-matching computation these signals are sorted according to information type (e.g., spatial vs. object vs. sensory-motor) and routed into appropriate subsequent processing streams (Griffiths & Warren, 2002).

An alternative view is that the PT is functionally segregated. This view is consistent with the fact that the PT contains a number of distinct cytoarchitectonic fields (Figure 6) (Galaburda & Sanides, 1980)<sup>5</sup>. The details of the functional organization of the PT remain to be specified but existing cytoarchitectonic data and new functional imaging data suggest a clear separation between spatial hearing-related functions of the one hand and sensory-motor functions (i.e., Spt) on the other. In functional MRI studies that map both speech related sensory-motor activity

and spatial hearing activity within the same subjects, these activations have been found to be distinct with sensory-motor function more posterior in the PT (Figure 10). This is consistent with human and comparative primate cortical anatomy. Although the PT is often thought of as part of auditory cortex, this is only true of the anterior portion. The posterior portion in macaque monkeys, referred to as area Tpt, is “an auditory-related multisensory area, but not part of auditory cortex” (p. 918) (Smiley et al., 2007), and in the human it has been characterized as “a transitional type of cortex between the specialized isocortices of the auditory region and the more generalized isocortex (integration cortex) of the inferior parietal lobule” (p. 609) (Galaburda & Sanides, 1980) (Figure 6).

All of the above evidence is consistent with the present conceptualization of area Spt in the posterior PT region as a (multi)sensory-motor integration area for vocal tract actions -- not just for speech functions - that is distinct from spatial hearing-related functions of the PT.

## A unifying framework: dual stream model

The processing of acoustic information in speech recognition and speech production involve partially overlapping, but also partially distinct neural circuits. Speech recognition relies primarily on neural circuits in the superior temporal lobes bilaterally, whereas speech production (and related processes such as verbal short-term memory) relies on a fronto-parietal/temporal circuit that is left hemisphere dominant. This divergence of processing streams is consistent with the fact that auditory/phonological information plays a role in (i) accessing lexical-semantic representations on the one hand and (ii) driving motor-speech articulation on the other. As lexical-semantic and motor-speech systems involve very different types of representations and processing mechanisms, it stands to reason that divergent pathways underlie the interface with auditory/phonological networks.

The dual interface requirements with respect to auditory/phonological processing is captured neuroanatomically by the Dual Stream model (Figure 1) (Hickok & Poeppel, 2000, 2004, 2007). The model is rooted in dual stream proposals in vision (Milner & Goodale, 1995) which distinguish between a ventral stream involved in visual object recognition (“what” stream) and dorsal stream involved in visual-motor integration (sometimes called a “how” stream). Accordingly, the Dual Stream model proposes that a ventral stream, which involves structures in the superior and middle portions of the temporal lobe, is involved in processing speech signals for comprehension (speech recognition), whereas a dorsal stream, which involves area Spt and posterior frontal lobe, is involved in translating speech signals into articulatory representations in the frontal lobe. The suggestion that the dorsal stream has an auditory-motor integration function differs from earlier arguments for a dorsal auditory “where” system (Rauschecker, 1998), but has gained support in recent years (Scott & Johnsrude, 2003; Warren et al., 2005; Wise et al., 2001). As indicated above, it is likely that a spatially-related processing system co-exists with, but is distinct from the sensory-motor integration system.

In contrast to the typical view that speech processing is mainly left hemisphere dependent, the model suggests that the ventral stream is bilaterally organized (although with important computational differences between the two hemispheres); thus, the ventral stream itself comprises parallel processing streams. This would explain the failure to find substantial speech recognition deficits following unilateral temporal lobe damage. The dorsal stream, on the other hand, is strongly left-dominant, explaining why production deficits are prominent sequelae of dorsal temporal and frontal lesions (Hickok & Poeppel, 2000, 2004, 2007).

---

<sup>5</sup>In fact, based on cytoarchitectonic data, it can be argued that it makes little sense to discuss the planum temporale as a functional region at all. Not only is the PT composed of several cytoarchitectonic fields, but these fields extend beyond the PT into both the parietal operculum and the lateral superior temporal gyrus and superior temporal sulcus (Galaburda & Sanides, 1980).

## Sentence- and grammatical-level functions

So far the discussion has been centered on phonemic and word-level processes. But what about higher-level mechanisms such as syntactic processing? This aspect of the neurology of language remains poorly understood, but some recent progress has been made.

For decades, Broca's region in the left inferior frontal lobe (see area 44 and 45 portions of blue shaded region in Figure 1) has been a centerpiece in hypotheses regarding the neural basis of syntactic processing both during production and comprehension. Broca's aphasics often exhibit agrammatic speech output, a tendency to omit grammatical function words and morphology (Friedmann, 2006; Goodglass et al., 2001; Hillis, 2007; Saffran, Schwartz, & Marin, 1980). In addition, such patients also tend to have difficulty comprehending syntactically complex sentences, specifically when structural relations between words in the sentence are required for accurate comprehension (Caramazza & Zurif, 1976; Grodzinsky, 1989; Hickok & Avrutin, 1996; Schwartz, Saffran, & Marin, 1980). For example, Broca's aphasics may have difficulty distinguishing between *He showed her the baby pictures* versus *He showed her baby the pictures*; the words are the same but because of their differing syntactic relations, the two sentences mean very different things. Because of the association between such "syntactic" deficits and Broca's aphasia, Broca's area was hypothesized to support some aspect of syntactic processing<sup>6</sup> (Caramazza & Berndt, 1978; Caramazza & Zurif, 1976; Friedmann, 2006; Grodzinsky, 1990, 2006).

The view that Broca's area played a role in syntactic functions was called into question, at least for receptive language, when it was demonstrated that agrammatic Broca's aphasics could judge the grammaticality of sentences (Linebarger, Schwartz, & Saffran, 1983) which clearly showed that syntactic abilities are to a large extent preserved in Broca's aphasia. Further, while functional imaging studies in neurologically healthy subjects showed that Broca's area responds more during the processing of syntactically complex compared to less complex sentences (Stromswold, Caplan, Alpert, & Rauch, 1996), other studies found that the response of Broca's region to language stimuli was not restricted to situations requiring syntactic analysis (Mazoyer et al., 1993). As a result, theoretical accounts of the role of Broca's region in syntactic processing emphasized either a very restricted syntactic deficit (Grodzinsky, 1986), processing deficits that disrupted the use of syntactic information in real time computation (Linebarger et al., 1983; Schwartz, Linebarger, Saffran, & Pate, 1987; Swinney, Zurif, & Nicol, 1989; Zurif, Swinney, Prather, Solomon, & Bushell, 1993), or working memory (Caplan & Waters, 1999; Martin, 2003). These issues are far from resolved (Grodzinsky & Santi, 2008), but one recent functional imaging study suggests that at least a portion of Broca's area supports syntactic processing via a working memory function. In this study, the previously observed sentence complexity effect (greater activation for syntactically more complex sentences relative to less complex sentences) was eliminated in the posterior portion of Broca's region (pars opercularis, ~BA 44) when subjects comprehended sentences while they concurrently rehearsed a set of syllables (Rogalsky, Matchin, & Hickok, 2008), but not when they performed a control concurrent finger tapping task (Figure 11). Rehearsing syllables has the effect of removing phonological working memory from the processing of such sentences (Baddeley, 1981), so the disappearance of the posterior Broca's region activation indicates that this region contributes to sentence processing via working memory. The role of the anterior region (pars triangularis), which showed a sentence complexity effect even during syllable rehearsal, remains to be determined.

---

<sup>6</sup>Damage to Broca's area alone is not sufficient to produce Broca's aphasia, rather a larger lesion is required (Mohr et al., 1978), but Broca's area is typically involved in the lesion of patients with Broca's aphasia (Damasio, 1992; Hillis, 2007).

Another brain region has recently emerged as a candidate for syntactic, or a least sentence-level processing: the lateral portion of the anterior temporal lobe (ATL) bilaterally (Figure 12). This region responds preferentially to structured sentences and less well to words in a foreign language, lists of nouns in one's own language, scrambled sentences, and meaningful environmental sound sequences (Friederici et al., 2000; Humphries et al., 2006; Humphries et al., 2005; Humphries et al., 2001; Mazoyer et al., 1993; Vandenberghe et al., 2002). The same region has been implicated in lesion studies of sentence comprehension: damage to the ATL is associated with increased difficulty comprehending syntactically complex sentences (Dronkers, Wilkins, Van Valin Jr., Redfern, & Jaeger, 2004) and in processing syntactically ambiguous sentences (Zaidel, Zaidel, Oxbury, & Oxbury, 1995). Some authors have suggested the ATL is specifically involved in performing syntactic computations (Humphries et al., 2006; Humphries et al., 2005), whereas others argue that it is involved in semantic integration within sentences (Vandenberghe et al., 2002), or combinatorial semantics, i.e., the process of integrating the meaning of words with their position in a hierarchical structure. These two possibilities are difficult to disentangle because combinatorial processes necessarily involve syntactic processes. A recent fMRI study attempted to tease apart these processes using a selective attention paradigm (Rogalsky & Hickok, 2009). First, the sentence-responsive ATL region was localized functionally via a standard sentence>word list contrast. Then subjects were asked either to monitor sentences for syntactic violations (*The flower are in the vase on the table*) or for combinatorial semantic "violations" (*The vase is in the flowers on the table*). Only responses to violation free sentences were analyzed. It was found that most of the ATL sentence responsive region was modulated by both attention tasks suggesting sensitivity to both syntactic and combinatorial semantic processes, although a subportion was only modulated by the semantic attention task (Rogalsky & Hickok, 2009).

Other investigations into the neural basis of grammatical functions have focused on a simpler linguistic forms such as morphological inflections (for reviews of the benefits of studying these forms see (Pinker, 1991,1999)). This work has suggested that disease involving frontal-basal ganglia circuits (e.g., in stroke, Parkinson's disease, Huntington's disease) produces disruptions to processing regularly inflected morphological forms (e.g., *walked, played*), which are argued to involve a grammatical computation, and less of an effect on irregular morphological forms (e.g., *run, slept*) which are argued to be stored in the mental lexicon and involve simple lexical access; disease affecting the lexical system in the temporal lobes (e.g., Alzheimer's, semantic dementia) produce the reverse pattern (Pinker & Ullman, 2002; Ullman, 2004; Ullman et al., 1997) (but see (Joanisse & Seidenberg, 2005)). One interpretation of this pattern is that grammatical processes are aligned with neural circuits supporting procedural memory whereas lexical processes are more aligned with circuits supporting declarative memory (Ullman, 2004).

As can be seen from this discussion, a number of different brain areas and circuits have been implicated in grammatical and sentence-level processing. This is perhaps not surprising given that processing higher-order aspects of language requires the integration of several sources of information including lexical-semantic, syntactic and prosodic/intonational cues among others. Processing this information in real time places additional demands associated with the need for timely lexical retrieval, ambiguity resolution (at multiple levels), and for integration of information over potentially long timescales (*It was [the driver] in the blue car that went through the red light who [almost hit me]*). These processing demands likely recruit a range of cognitive processes such as working memory, attention, sequence processing, and response selection/inhibition. Investigators are now beginning to sort through some of these contributors to grammatical/sentence level processing (Caplan, Alpert, Waters, & Olivieri, 2000; Fiebach & Schubotz, 2006; Grodzinsky & Santi, 2008; Novick, Trueswell, & Thompson-Schill, 2005; Rogalsky, Matchin et al., 2008; Schubotz & von Cramon, 2004), but at present it remains

unclear how much of the “sentence processing network” reflects these processes as opposed to linguistic specific grammatical computations.

## Summary and conclusions

The transformation between acoustic wave form and thought is complex operation that is achieved via multiple interacting brain systems. Much remains to be understood particularly in terms of the neural circuits involved in higher-level aspects of language processing (sentence-level and grammatical processes) and in terms of the neural computations that are carried out at all stages of processing. Aspects of the functional architecture of word-level circuits are beginning to come into focus, however. Early stages of speech recognition are bilaterally (but not necessarily symmetrically) organized in the superior temporal lobe. Beyond that the speech processing stream diverges into two pathways, one that interfaces auditory-phonological information with the conceptual system and therefore supports comprehension, and another that interfaces auditory-phonological information with the motor system and therefore support sensory-motor integration functions such as aspects of speech production and phonological short-term memory. The sensory-motor interface is not speech specific. The sensory-motor interface may be the pathway by which motor knowledge of speech can provide top-down modulation of speech perception, but there is strong evidence *against* the view that motor or sensory-motor circuits play a central role in speech recognition. Processing of grammatical information likely involves a number of brain circuits that may include portions of Broca’s region, fronto-basal ganglia circuits, and the anterior temporal lobe; it is unclear to what extent these networks are language specific.

## Acknowledgments

Supported by NIH grants R01 DC03681 and R01 DC009659

## References

- Aboitiz F, García VR. The evolutionary origin of language areas in the human brain. A neuroanatomical perspective. *Brain Research Reviews* 1997;25:381–396. [PubMed: 9495565]
- Abrams DA, Nicol T, Zecker S, Kraus N. Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J Neurosci* 2008;28(15):3958–3965. [PubMed: 18400895]
- Andersen R. Multimodal integration for the representation of space in the posterior parietal cortex. *Philos Trans R Soc Lond B Biol Sci* 1997;352:1421–1428. [PubMed: 9368930]
- Anderson JM, Gilmore R, Roper S, Crosson B, Bauer RM, Nadeau S, Beversdorf DQ, Cibula J, Rogish M III, Kortencamp S, Hughes JD, Gonzalez Rothi LJ, Heilman KM. Conduction aphasia and the arcuate fasciculus: A reexamination of the Wernicke-Geschwind model. *Brain and Language* 1999;70:1–12. [PubMed: 10534369]
- Bachman, DL.; Albert, ML. Auditory comprehension in aphasia. In: Boller, F.; Grafman, J., editors. *Handbook of neuropsychology*. Vol. 1. New York: Elsevier; 1988. p. 281–306.
- Baddeley A. The role of subvocalisation in reading. *Q J Exp Psychol* 1981;33A:439–454.
- Baddeley A, Gathercole S, Papagno C. The phonological loop as a language learning device. *Psychol Rev* 1998;105(1):158–173. [PubMed: 9450375]
- Baddeley AD. Working memory. *Science* 1992;255:556–559. [PubMed: 1736359]
- Baker E, Blumstein SE, Goodglass H. Interaction between phonological and semantic factors in auditory comprehension. *Neuropsychologia* 1981;19:1–15. [PubMed: 7231654]
- Baldo JV, Klostermann EC, Dronkers NF. It’s either a cook or a baker: patients with conduction aphasia get the gist but lose the trace. *Brain Lang* 2008;105(2):134–140. [PubMed: 18243294]
- Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, Dronkers NF. Voxel-based lesion-symptom mapping. *Nat Neurosci* 2003;6(5):448–450. [PubMed: 12704393]

- Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Posing ET. Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex* 2000;10:512–528. [PubMed: 10847601]
- Binder JR, Frost JA, Hammeke TA, Cox RW, Rao SM, Prieto T. Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience* 1997;17:353–362. [PubMed: 8987760]
- Binder JR, Rao SM, Hammeke TA, Yetkin FZ, Jesmanowicz A, Bandettini PA, Wong EC, Estkowski LD, Goldstein MD, Haughton VM, Hyde JS. Functional magnetic resonance imaging of human auditory cortex. *Annals of Neurology* 1994;35:662–672. [PubMed: 8210222]
- Binder JT, Frost JA, Hammeke TA, Rao SM, Cox RW. Function of the left planum temporale in auditory and linguistic processing. *Brain* 1996;119:1239–1247. [PubMed: 8813286]
- Bock, K. Language production. In: Wilson, RA.; Keil, FC., editors. *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press; 1999. p. 453-456.
- Boemio A, Fromm S, Braun A, Poeppel D. Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat Neurosci* 2005;8(3):389–395. [PubMed: 15723061]
- Boller F, Marcie P. Possible role of abnormal auditory feedback in conduction aphasia. *Neuropsychologia* 1978;16(4):521–524. [PubMed: 692865]
- Boller F, Vrtunski PB, Kim Y, Mack JL. Delayed auditory feedback and aphasia. *Cortex* 1978;14(2): 212–226. [PubMed: 679702]
- Breese EL, Hillis AE. Auditory comprehension: is multiple choice really good enough? *Brain Lang* 2004;89(1):3–8. [PubMed: 15010231]
- Buchman AS, Garron DC, Trost-Cardamone JE, Wichter MD, Schwartz M. Word deafness: One hundred years later. *Journal of Neurology, Neurosurgery, and Psychiatry* 1986;49:489–499.
- Buchsbaum B, Hickok G, Humphries C. Role of Left Posterior Superior Temporal Gyrus in Phonological Processing for Speech Perception and Production. *Cognitive Science* 2001;25:663–678.
- Buchsbaum BR, D’Esposito M. The search for the phonological store: from loop to convolution. *J Cogn Neurosci* 2008;20(5):762–778. [PubMed: 18201133]
- Buchsbaum BR, Olsen RK, Koch P, Berman KF. Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron* 2005;48(4):687–697. [PubMed: 16301183]
- Buchsbaum BR, Olsen RK, Koch PF, Kohn P, Kippenhan JS, Berman KF. Reading, hearing, and the planum temporale. *Neuroimage* 2005;24(2):444–454. [PubMed: 15627586]
- Burnett TA, Senner JE, Larson CR. Voice F0 responses to pitch-shifted auditory feedback: a preliminary study. *J Voice* 1997;11(2):202–211. [PubMed: 9181544]
- Calvert GA, Bullmore ET, Brammer MJ, Campbell R, Williams SCR, McGuire PK, Woodruff PWR, Iversen SD, David AS. Activation of auditory cortex during silent lipreading. *Science* 1997;276:593–596. [PubMed: 9110978]
- Calvert GA, Campbell R. Reading speech from still and moving faces: The neural substrates of visible speech. *Journal of Cognitive Neuroscience* 2003;15:57–70. [PubMed: 12590843]
- Caplan D, Alpert N, Waters G, Olivieri A. Activation of Broca’s area by syntactic processing under conditions of concurrent articulation. *Hum Brain Mapp* 2000;9(2):65–71. [PubMed: 10680763]
- Caplan D, Waters G. Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences* 1999;22:114–126.
- Caramazza A, Berndt RS. Semantic and syntactic processes in aphasia: a review of the literature. *Psychol Bull* 1978;85(4):898–918. [PubMed: 356076]
- Caramazza A, Mahon BZ. The organization of conceptual knowledge: the evidence from category-specific semantic deficits. *Trends Cogn Sci* 2003;7(8):354–361. [PubMed: 12907231]
- Caramazza A, Zurif EB. Dissociation of algorithmic and heuristic processes in sentence comprehension: Evidence from aphasia. *Brain and Language* 1976;3:572–582. [PubMed: 974731]
- Chertkow H, Bub D, Deaudon C, Whitehead V. On the status of object concepts in aphasia. *Brain Lang* 1997;58(2):203–232. [PubMed: 9182748]

- Christen HJ, Hanefeld F, Kruse E, Imhauser S, Ernst JP, Finkenstaedt M. Foix-Chavany-Marie (anterior operculum) syndrome in childhood: a reappraisal of Worster-Drought syndrome. *Dev Med Child Neurol* 2000;42(2):122–132. [PubMed: 10698330]
- Colby CL, Goldberg ME. Space and attention in parietal cortex. *Annual Review of Neuroscience* 1999;22:319–349.
- Craigheo L, Metta G, Sandini G, Fadiga L. The mirror-neurons system: data and models. *Prog Brain Res* 2007;164:39–59. [PubMed: 17920425]
- D'Ausilio A, Pulvermuller F, Salmas P, Bufalari I, Begliomini C, Fadiga L. The motor somatotopy of speech perception. *Curr Biol* 2009;19(5):381–385. [PubMed: 19217297]
- D'Ausilio A, Pulvermuller F, Salmas P, Bufalari I, Begliomini C, Fadiga L. The motor somatotopy of speech perception. *Curr Biol*. 2009
- Damasio AR. The brain binds entities and events by multiregional activation from convergence zones. *Neural Computation* 1989;1:123–132.
- Damasio, AR. Signs of aphasia. In: Sarno, MT., editor. *Acquired aphasia*. 2. San Diego: Academic Press; 1991. p. 27-43.
- Damasio AR. Aphasia. *New England Journal of Medicine* 1992;326:531–539. [PubMed: 1732792]
- Damasio, H. Neuroanatomical correlates of the aphasias. In: Sarno, M., editor. *Acquired aphasia*. 2. San Diego: Academic Press; 1991. p. 45-71.
- Damasio H, Damasio AR. The anatomical basis of conduction aphasia. *Brain* 1980;103:337–350. [PubMed: 7397481]
- Dell, GS. Speaking and misspeaking. In: Gietman, LR.; Liberman, M., editors. *An invitation to cognitive science: Language*. 2. Vol. 1. Cambridge, MA: MIT Press; 1995. p. 183-208.
- Dell GS, Schwartz MF, Martin N, Saffran EM, Gagnon DA. Lexical access in aphasic and nonaphasic speakers. *Psychological Review* 1997;104:801–838. [PubMed: 9337631]
- di Pellegrino G, Fadiga L, Fogassi L, Gallese V, Rizzolatti G. Understanding motor events: a neurophysiological study. *Exp Brain Res* 1992;91(1):176–180. [PubMed: 1301372]
- Doupe AJ, Kuhl PK. Birdsong and human speech: Common themes and mechanisms. *Annual Review of Neuroscience* 1999;22:567–631.
- Dronkers, NF.; Redfern, BB.; Knight, RT. The neural architecture of language disorders. In: Gazzaniga, MS., editor. *The new cognitive neurosciences*. Cambridge, MA: MIT Press; 2000. p. 949-958.
- Dronkers, NF.; Wilkins, DP.; Van Valin, RD., Jr; Redfern, BB.; Jaeger, JJ. Lesion analysis of brain regions involved in language comprehension. In: Hickok, G.; Poeppel, D., editors. *The New Functional Anatomy of Language: A special issue of Cognition*. 2004.
- Eimas PD, Siqueland ER, Jusczyk P, Vigorito J. Speech perception in infants. *Science* 1971;171(968):303–306. [PubMed: 5538846]
- Fadiga L, Craighero L. Hand actions and speech representation in Broca's area. *Cortex* 2006;42(4):486–490. [PubMed: 16881255]
- Fadiga L, Craighero L, Buccino G, Rizzolatti G. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur J Neurosci* 2002;15(2):399–402. [PubMed: 11849307]
- Fiebach CJ, Schubotz RI. Dynamic anticipatory processing of hierarchical sequential events: a common role for Broca's area and ventral premotor cortex across domains? *Cortex* 2006;42(4):499–502. [PubMed: 16881258]
- Foundas AL, Bollich AM, Feldman J, Corey DM, Hurley M, Lemen LC, Heilman KM. Aberrant auditory processing and atypical planum temporale in developmental stuttering. *Neurology* 2004;63(9):1640–1646. [PubMed: 15534249]
- Friederici AD, Meyer M, von Cramon DY. Auditory language comprehension: An event-related fMRI study on the processing of syntactic and lexical information. *Brain and Language* 2000;74:289–300. [PubMed: 10950920]
- Friedmann, N. Speech production in Broca's agrammatic aphasia: syntactic tree pruning. In: Grodzinsky, Y.; Amunts, K., editors. *Broca's region*. Oxford: Oxford University Press; 2006. p. 63-82.
- Fromkin V. The non-anomalous nature of anomalous utterances. *Language* 1971;47:27–52.
- Fuster, JM. *Memory in the cerebral cortex*. Cambridge, MA: MIT Press; 1995.

- Gainotti G, Micelli G, Silveri MC, Villa G. Some anatomo-clinical aspects of phonemic and semantic comprehension disorders in aphasia. *Acta Neurologica Scandinavica* 1982;66:652–665. [PubMed: 6187173]
- Galaburda A, Sanides F. Cytoarchitectonic organization of the human auditory cortex. *Journal of Comparative Neurology* 1980;190:597–610. [PubMed: 6771305]
- Galantucci B, Fowler CA, Turvey MT. The motor theory of speech perception reviewed. *Psychon Bull Rev* 2006;13(3):361–377. [PubMed: 17048719]
- Gallese V, Fadiga L, Fogassi L, Rizzolatti G. Action recognition in the premotor cortex. *Brain* 1996;119 (Pt 2):593–609. [PubMed: 8800951]
- Gannon PJ, Holloway RL, Broadfield DC, Braun AR. Asymmetry of the chimpanzee planum temporale: humanlike pattern of Wernicke's brain language area homolog. *Science* 1998;279:220–222. [PubMed: 9422693]
- Geschwind N. Disconnexion syndromes in animals and man. *Brain* 1965;88:237–294. 585–644. [PubMed: 5318481]
- Geschwind N. Aphasia. *New England Journal of Medicine* 1971;284:654–656. [PubMed: 5545606]
- Geschwind N, Levitsky W. Human brain: Left-right asymmetries in temporal speech region. *Science* 1968;161:186–187. [PubMed: 5657070]
- Giraud AL, Kleinschmidt A, Poeppel D, Lund TE, Frackowiak RS, Laufs H. Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 2007;56(6):1127–1134. [PubMed: 18093532]
- Goodglass, H. Diagnosis of conduction aphasia. In: Kohn, SE., editor. *Conduction aphasia*. Hillsdale, N.J: Lawrence Erlbaum Associates; 1992. p. 39-49.
- Goodglass, H. *Understanding aphasia*. San Diego: Academic Press; 1993.
- Goodglass, H.; Kaplan, E.; Barresi, B. *The assessment of aphasia and related disorders*. 3. Philadelphia: Lippincott Williams & Wilkins; 2001.
- Gorno-Tempini ML, Dronkers NF, Rankin KP, Ogar JM, Phengrasamy L, Rosen HJ, Johnson JK, Weiner MW, Miller BL. Cognition and anatomy in three variants of primary progressive aphasia. *Ann Neurol* 2004;55(3):335–346. [PubMed: 14991811]
- Graves WW, Grabowski TJ, Mahta S, Gordon JK. A neural signature of phonological access: distinguishing the effects of word frequency from familiarity and length in overt picture naming. *Journal of Cognitive Neuroscience* 2007;19:617–631. [PubMed: 17381253]
- Grefkes C, Fink GR. The functional organization of the intraparietal sulcus in humans and monkeys. *J Anat* 2005;207(1):3–17. [PubMed: 16011542]
- Griffiths TD, Warren JD. The planum temporale as a computational hub. *Trends Neurosci* 2002;25(7):348–353. [PubMed: 12079762]
- Grodzinsky Y. Language deficits and the theory of syntax. *Brain and Language* 1986;27:135–159. [PubMed: 3947938]
- Grodzinsky Y. Agrammatic comprehension of relative clauses. *Brain and Language* 1989;31:480–499. [PubMed: 2478254]
- Grodzinsky, Y. *Theoretical perspectives on language deficits*. Cambridge, MA: MIT Press; 1990.
- Grodzinsky, Y. A blueprint for a brain map of syntax. In: Grodzinsky, Y.; Amunts, K., editors. *Broca's region*. Oxford: Oxford University Press; 2006. p. 83-107.
- Grodzinsky Y, Santi A. The battle for Broca's region. *Trends Cogn Sci* 2008;12(12):474–480. [PubMed: 18930695]
- Hart JJ, Gordon B. Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Annals of Neurology* 1990;27:226–231. [PubMed: 2327733]
- Hickok, G. Speech perception, conduction aphasia, and the functional neuroanatomy of language. In: Grodzinsky, Y.; Shapiro, L.; Swinney, D., editors. *Language and the brain*. San Diego: Academic Press; 2000. p. 87-104.
- Hickok G, Avrutin S. Comprehension of wh-questions in two Broca's aphasics. *Brain and Language* 1996;52:314–327. [PubMed: 8811962]

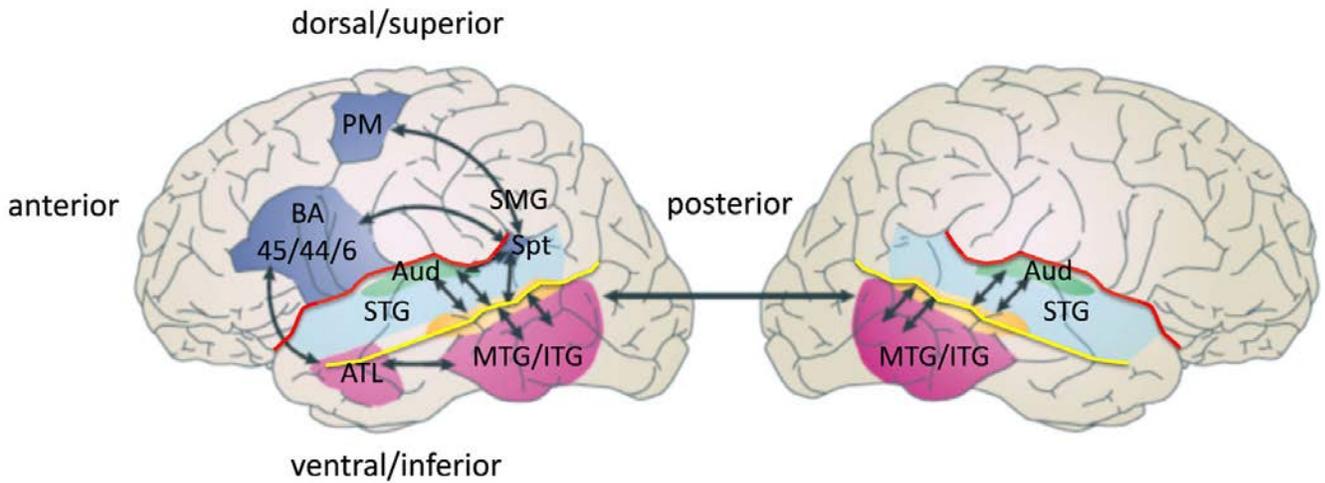
- Hickok G, Buchsbaum B, Humphries C, Muftuler T. Auditory-motor interaction revealed by fMRI: Speech, music, and working memory in area *Spt*. *Journal of Cognitive Neuroscience* 2003;15:673–682. [PubMed: 12965041]
- Hickok G, Erhard P, Kassubek J, Helms-Tillery AK, Naeve-Velguth S, Strupp JP, Strick PL, Ugurbil K. A functional magnetic resonance imaging study of the role of left posterior superior temporal gyrus in speech production: implications for the explanation of conduction aphasia. *Neuroscience Letters* 2000;287:156–160. [PubMed: 10854735]
- Hickok G, Okada K, Barr W, Pa J, Rogalsky C, Donnelly K, Barde L, Grant A. Bilateral capacity for speech sound processing in auditory comprehension: evidence from Wada procedures. *Brain Lang* 2008;107(3):179–184. [PubMed: 18976806]
- Hickok G, Okada K, Serences JT. Area Spt in the Human Planum Temporale Supports Sensory-Motor Integration for Speech Processing. *J Neurophysiol.* 2009
- Hickok G, Poeppel D. Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences* 2000;4:131–138. [PubMed: 10740277]
- Hickok G, Poeppel D. Dorsal and ventral streams: A framework for understanding aspects of the functional anatomy of language. *Cognition* 2004;92:67–99. [PubMed: 15037127]
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci* 2007;8(5):393–402. [PubMed: 17431404]
- Hillis AE. Aphasia: progress in the last quarter of a century. *Neurology* 2007;69(2):200–213. [PubMed: 17620554]
- Hodges JR, Patterson K. Semantic dementia: a unique clinicopathological syndrome. *Lancet Neurol* 2007;6(11):1004–1014. [PubMed: 17945154]
- Houde JF, Jordan MI. Sensorimotor adaptation in speech production. *Science* 1998;279:1213–1216. [PubMed: 9469813]
- Humphries C, Binder JR, Medler DA, Liebenthal E. Syntactic and semantic modulation of neural activity during auditory sentence comprehension. *J Cogn Neurosci* 2006;18(4):665–679. [PubMed: 16768368]
- Humphries C, Love T, Swinney D, Hickok G. Response of anterior temporal cortex to syntactic and prosodic manipulations during sentence processing. *Human Brain Mapping* 2005;26:128–138. [PubMed: 15895428]
- Humphries C, Willard K, Buchsbaum B, Hickok G. Role of anterior temporal cortex in auditory sentence comprehension: An fMRI study. *Neuroreport* 2001;12:1749–1752. [PubMed: 11409752]
- Indefrey P, Levelt WJ. The spatial and temporal signatures of word production components. *Cognition* 2004;92(1–2):101–144. [PubMed: 15037128]
- Indefrey, P.; Levelt, WJM. The neural correlates of language production. In: Gazzaniga, MS., editor. *The new cognitive neurosciences*. Cambridge, MA: MIT Press; 2000. p. 845-865.
- Jacquemot C, Scott SK. What is the relationship between phonological short-term memory and speech processing? *Trends Cogn Sci* 2006;10:480–486. [PubMed: 16997610]
- Joanisse MF, Seidenberg MS. Imaging the past: neural activation in frontal and temporal regions during regular and irregular past-tense processing. *Cogn Affect Behav Neurosci* 2005;5(3):282–296. [PubMed: 16396090]
- Jones DM, Hughes RW, Macken WJ. The phonological store abandoned. *QJ Exp Psychol (Colchester)* 2007;60(4):505–511.
- Jones DM, Macken WJ. Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 1996;19:369–381.
- Jones DM, Macken WJ, Nicholls AP. The phonological store of working memory: is it phonological and is it a store? *J Exp Psychol Learn Mem Cogn* 2004;30(3):656–674. [PubMed: 15099134]
- Kuhl PK, Miller JD. Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants. *Science* 1975;190:69–72. [PubMed: 1166301]
- Lenneberg EH. Understanding language without ability to speak: a case report. *Journal of Abnormal and Social Psychology* 1962;65:419–425. [PubMed: 13929636]
- Levelt, WJM. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press; 1989.

- Levelt WJM, Praamstra P, Meyer AS, Helenius P, Salmelin R. An MEG study of picture naming. *Journal of Cognitive Neuroscience* 1998;10:553–567. [PubMed: 9802989]
- Levine DN, Mohr JP. Language after bilateral cerebral infarctions: role of the minor hemisphere in speech. *Neurology* 1979;29(7):927–938. [PubMed: 572942]
- Lieberman AM. Some results of research on speech perception. *Journal of the Acoustical Society of America* 1957;29(1):117–123.
- Lieberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. *Psychol Rev* 1967;74(6):431–461. [PubMed: 4170865]
- Lieberman AM, Mattingly IG. The motor theory of speech perception revised. *Cognition* 1985;21:1–36. [PubMed: 4075760]
- Lichtheim L. On aphasia. *Brain* 1885;7:433–484.
- Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. Neural substrates of phonemic perception. *Cereb Cortex* 2005;15(10):1621–1631. [PubMed: 15703256]
- Linebarger MC, Schwartz M, Saffran E. Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition* 1983;13:361–393. [PubMed: 6683142]
- Lotto AJ, Hickok GS, Holt LL. Reflections on Mirror Neurons and Speech Perception. *Trends Cogn Sci* 2009;13:110–114. [PubMed: 19223222]
- Luce PA, Pisoni DB. Recognizing spoken words: The neighborhood activation model. *Ear and Hearing* 1998;19:1–36. [PubMed: 9504270]
- Luo H, Poeppel D. Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 2007;54(6):1001–1010. [PubMed: 17582338]
- Marslen-Wilson WD. Functional parallelism in spoken word-recognition. *Cognition* 1987;25:71–102. [PubMed: 3581730]
- Martin, A. The organization of semantic knowledge and the origin of words in the brain. In: Jablonski, NG.; Aiello, LC., editors. *The origins and diversification of language*. San Francisco: California Academy of Sciences; 1998. p. 69–88.
- Martin A, Chao LL. Semantic memory and the brain: structure and processes. *Curr Opin Neurobiol* 2001;11(2):194–201. [PubMed: 11301239]
- Martin RC. Language processing: functional organization and neuroanatomical basis. *Annu Rev Psychol* 2003;54:55–89. [PubMed: 12359917]
- Mazoyer BM, Tzourio N, Frak V, Syrota A, Murayama N, Levrier O, Salamon G, Dehaene S, Cohen L, Mehler J. The cortical representation of *speech*. *Journal of Cognitive Neuroscience* 1993;5:467–479.
- McClelland JL, Elman JL. The TRACE model of speech perception. *Cognitive Psychology* 1986;18:1–86. [PubMed: 3753912]
- McGlone J. Speech comprehension after unilateral injection of sodium amytal. *Brain and Language* 1984;22:150–157. [PubMed: 6722523]
- McNealy K, Mazziotta JC, Dapretto M. Cracking the language code: neural mechanisms underlying speech parsing. *J Neurosci* 2006;26(29):7629–7639. [PubMed: 16855090]
- Miceli G, Gainotti G, Caltagirone C, Masullo C. Some aspects of phonological impairment in aphasia. *Brain and Language* 1980;11:159–169. [PubMed: 7427715]
- Miglioretti DL, Boatman D. Modeling variability in cortical representations of human complex sound perception. *Exp Brain Res* 2003;153(3):382–387. [PubMed: 14534769]
- Milner, AD.; Goodale, MA. *The visual brain in action*. Oxford: Oxford University Press; 1995.
- Mohr JP, Pessin MS, Finkelstein S, Funkenstein HH, Duncan GW, Davis KR. Broca's aphasia: Pathological and clinical. *Neurology* 1978;28:311–324. [PubMed: 565019]
- Naeser MA, Palumbo CL, Helm-Estabrooks N, Stiassny-Eder D, Albert ML. Severe nonfluency in aphasia: Role of the medial subcallosal fasciculus and other white matter pathways in recovery of spontaneous speech. *Brain* 1989;112:1–38. [PubMed: 2917272]
- Narain C, Scott SK, Wise RJ, Rosen S, Leff A, Iversen SD, Matthews PM. Defining a left-lateralized response specific to intelligible speech using fMRI. *Cereb Cortex* 2003;13(12):1362–1368. [PubMed: 14615301]

- Novick JM, Trueswell JC, Thompson-Schill SL. Cognitive control and parsing: reexamining the role of Broca's area in sentence comprehension. *Cogn Affect Behav Neurosci* 2005;5(3):263–281. [PubMed: 16396089]
- Obleser J, Zimmermann J, Van Meter J, Rauschecker JP. Multiple Stages of Auditory Speech Perception Reflected in Event-Related fMRI. *Cereb Cortex*. 2006
- Okada K, Hickok G. Identification of lexical-phonological networks in the superior temporal sulcus using fMRI. *Neuroreport* 2006;17:1293–1296. [PubMed: 16951572]
- Okada K, Hickok G. Two cortical mechanisms support the integration of visual and auditory speech: A hypothesis and preliminary data. *Neurosci Lett* 2009;452(3):219–223. [PubMed: 19348727]
- Okada K, Smith KR, Humphries C, Hickok G. Word Length Modulates Neural Activity in Auditory Cortex During Covert Object Naming. *Neuroreport* 2003;14:2323–2326. [PubMed: 14663184]
- Pa J, Hickok G. A parietal-temporal sensory-motor integration area for the human vocal tract: Evidence from an fMRI study of skilled musicians. *Neuropsychologia* 2008;46:362–368. [PubMed: 17709121]
- Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat Rev Neurosci* 2007;8(12):976–987. [PubMed: 18026167]
- Paulesu E, Vallar G, Berlinger M, Signorini M, Vitali P, Burani C, Perani D, Fazio F. Supercalifragilisticexpialidocious: how the brain learns words never heard before. *Neuroimage* 2009;45(4):1368–1377. [PubMed: 19171195]
- Pinker S. Rules of language. *Science* 1991;253:530–535. [PubMed: 1857983]
- Pinker, S. Words and rules: The ingredients of language. New York: Basic Books; 1999.
- Pinker S, Ullman MT. The past and future of the past tense. *Trends Cogn Sci* 2002;6(11):456–463. [PubMed: 12457895]
- Poeppl D. The analysis of speech in different temporal integration windows: cerebral lateralization as “asymmetric sampling in time”. *Speech Communication* 2003;41:245–255.
- Price CJ, Wise RJS, Warburton EA, Moore CJ, Howard D, Patterson K, Frackowiak RSJ, Friston KJ. Hearing and saying: The functional neuro-anatomy of auditory word processing. *Brain* 1996;119:919–931. [PubMed: 8673502]
- Pulvermuller F. Hebb's concept of cell assemblies and the psychophysiology of word processing. *Psychophysiology* 1996;33(4):317–333. [PubMed: 8753932]
- Rauschecker JP. Cortical processing of complex sounds. *Current Opinion in Neurobiology* 1998;8(4): 516–521. [PubMed: 9751652]
- Remez RE, Rubin PE, Pisoni DB, Carrell TD. Speech perception without traditional speech cues. *Science* 1981;212:947–950. [PubMed: 7233191]
- Rissman J, Eliassen JC, Blumstein SE. An event-related FMRI investigation of implicit semantic priming. *J Cogn Neurosci* 2003;15(8):1160–1175. [PubMed: 14709234]
- Rizzolatti G, Arbib M. Language within our grasp. *Trends in Neurosciences* 1998;21:188–194. [PubMed: 9610880]
- Rizzolatti G, Fogassi L, Gallese V. Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci* 2001;2(9):661–670. [PubMed: 11533734]
- Rodd JM, Davis MH, Johnsrude IS. The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex* 2005;15:1261–1269. [PubMed: 15635062]
- Rogalsky C, Hickok G. Selective Attention to Semantic and Syntactic Features Modulates Sentence Processing Networks in Anterior Temporal Cortex. *Cereb Cortex* 2009;19:786–796. [PubMed: 18669589]
- Rogalsky C, Matchin W, Hickok G. Broca's Area, Sentence Comprehension, and Working Memory: An fMRI Study. *Front Hum Neurosci* 2008;2:14. [PubMed: 18958214]
- Rogalsky C, Pitz E, Hillis AE, Hickok G. Auditory word comprehension impairment in acute stroke: relative contribution of phonemic versus semantic factors. *Brain Lang* 2008;107(2):167–169. [PubMed: 18823655]
- Ruchkin DS, Grafman J, Cameron K, Berndt RS. Working memory retention systems: a state of activated long-term memory. *Behav Brain Sci* 2003;26:709–777. [PubMed: 15377128]

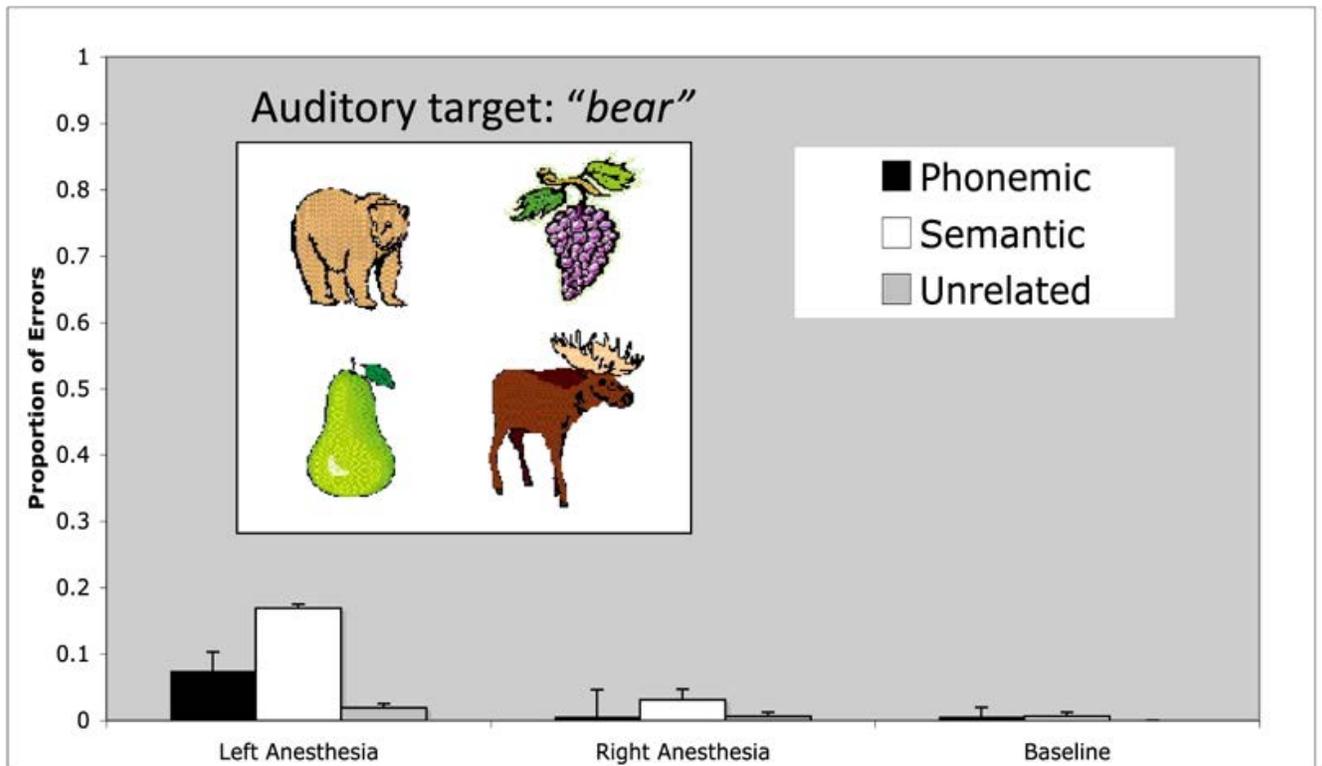
- Saffran EM, Schwartz MF, Marin OS. The word order problem in agrammatism. II. Production. *Brain Lang* 1980;10(2):263–280. [PubMed: 7407547]
- Schlaug G, Jancke L, Huang Y, Steinmetz H. In vivo evidence of structural brain asymmetry in musicians. *Science* 1995;267:699–701. [PubMed: 7839149]
- Schlosser MJ, Aoyagi N, Fulbright RK, Gore JC, McCarthy G. Functional MRI studies of auditory comprehension. *Human Brain Mapping* 1998;6:1–13. [PubMed: 9673659]
- Schubotz RI, von Cramon DY. Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. *J Neurosci* 2004;24(24):5467–5474. [PubMed: 15201318]
- Schwartz MF, Linebarger MC, Saffran EM, Pate DS. Syntactic transparency and sentence interpretation in aphasia. *Language and Cognitive Processes* 1987;2:85–113.
- Schwartz MF, Saffran EM, Marin OS. The word order problem in agrammatism. I. Comprehension. *Brain Lang* 1980;10(2):249–262. [PubMed: 7407546]
- Scott SK, Blank CC, Rosen S, Wise RJS. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 2000;123:2400–2406. [PubMed: 11099443]
- Scott SK, Johnsrude IS. The neuroanatomical and functional organization of speech perception. *Trends Neurosci* 2003;26(2):100–107. [PubMed: 12536133]
- Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. *Science* 1995;270:303–304. [PubMed: 7569981]
- Shelton JR, Caramazza A. Deficits in lexical and semantic processing: Implications for models of normal language. *Psychonomic Bulletin & Review* 1999;6:5–27. [PubMed: 12199314]
- Smiley JF, Hackett TA, Ulbert I, Karmas G, Lakatos P, Javitt DC, Schroeder CE. Multisensory convergence in auditory cortex. I. Cortical connections of the caudal superior temporal plane in macaque monkeys. *J Comp Neurol* 2007;502(6):894–923. [PubMed: 17447261]
- Smith KR, Okada K, Saberi K, Hickok G. Human cortical motion areas are not motion selective. *Neuroreport* 2004;9:1523–1526. [PubMed: 15194888]
- Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJ. Converging language streams in the human temporal lobe. *J Neurosci* 2006;26(28):7328–7336. [PubMed: 16837579]
- Stevens KN. Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustic Society of America* 2002;111:1872–1891.
- Stromswold K, Caplan D, Alpert N, Rauch S. Localization of syntactic comprehension by positron emission tomography. *Brain Lang* 1996;52(3):452–473. [PubMed: 8653390]
- Stuart A, Frazier CL, Kalinowski J, Vos PW. The effect of frequency altered feedback on stuttering duration and type. *J Speech Lang Hear Res* 2008;51(4):889–897. [PubMed: 18658059]
- Stuart A, Kalinowski J, Rastatter MP, Lynch K. Effect of delayed auditory feedback on normal speakers at two speech rates. *J Acoust Soc Am* 2002;111(5 Pt 1):2237–2241. [PubMed: 12051443]
- Sussman HM. Representation of phonological categories: a functional role for auditory columns. *Brain Lang* 2002;80(1):1–13. [PubMed: 11817886]
- Swinney D, Zurif E, Nicol J. The effects of focal brain damage on sentence processing: An examination of the neurological organization of a mental module. *Journal of Cognitive Neuroscience* 1989;1:25–37.
- Terao Y, Ugawa Y, Yamamoto T, Sakurai Y, Masumoto T, Abe O, Masutani Y, Aoki S, Tsuji S. Primary face motor area as the motor representation of articulation. *J Neurol* 2007;254(4):442–447. [PubMed: 17380243]
- Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 2008;39(3):1429–1443. [PubMed: 18035557]
- Ullman MT. Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 2004;92(1–2):231–270. [PubMed: 15037131]
- Ullman MT, Corkin S, Coppola M, Hickok G, Growdon JH, Koroshetz WJ, Pinker S. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedure *system*. *Journal of Cognitive Neuroscience* 1997;9:266–276.
- Ungerleider, LG.; Mishkin, M. Two cortical visual systems. In: Ingle, DJ.; Goodale, MA.; Mansfield, RJW., editors. *Analysis of visual behavior*. Cambridge, MA: MIT Press; 1982. p. 549–586.

- Vandenberghe R, Nobre AC, Price CJ. The response of left temporal cortex to sentences. *J Cogn Neurosci* 2002;14(4):550–560. [PubMed: 12126497]
- Vouloumanos A, Kiehl KA, Werker JF, Liddle PF. Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech. *J Cogn Neurosci* 2001;13(7):994–1005. [PubMed: 11595101]
- Wada J, Rasmussen T. Intracarotid injection of sodium amytal for the lateralization of cerebral speech dominance. *Journal of Neurosurgery* 1960;17:266–282.
- Waldstein RS. Effects of postlingual deafness on speech production: Implications for the role of auditory feedback. *Journal of the Acoustical Society of America* 1989;88:2099–2144. [PubMed: 2269726]
- Warren JD, Griffiths TD. Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. *The Journal of Neuroscience* 2003;23:5799–5804. [PubMed: 12843284]
- Warren JD, Zielinski BA, Green GG, Rauschecker JP, Griffiths TD. Perception of sound-source motion by the human brain. *Neuron* 2002;34(1):139–148. [PubMed: 11931748]
- Warren JE, Wise RJ, Warren JD. Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends Neurosci* 2005;28(12):636–643. [PubMed: 16216346]
- Weller M. Anterior opercular cortex lesions cause dissociated lower cranial nerve palsies and anarthria but no aphasia: Foix-Chavany-Marie syndrome and “automatic voluntary dissociation” revisited. *J Neurol* 1993;240(4):199–208. [PubMed: 7684439]
- Wernicke, C. The symptom complex of aphasia: A psychological study on an anatomical basis. In: Cohen, RS.; Wartofsky, MW., editors. *Boston studies in the philosophy of science*. Dordrecht: D. Reidel Publishing Company; 18741969. p. 34-97.
- Wise RJS, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA. Separate neural sub-systems within “Wernicke’s area”. *Brain* 2001;124:83–95. [PubMed: 11133789]
- Yates AJ. Delayed auditory feedback. *Psychological Bulletin* 1963;60:213–251. [PubMed: 14002534]
- Zaidel DW, Zaidel E, Oxbury SM, Oxbury JM. The interpretation of sentence ambiguity in patients with unilateral focal brain surgery. *Brain Lang* 1995;51(3):458–468. [PubMed: 8719077]
- Zaidel, E. Language in the right hemisphere. In: Benson, DF.; Zaidel, E., editors. *The dual brain: Hemispheric specialization in humans*. New York: Guilford Press; 1985. p. 205-231.
- Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. *Trends in Cognitive Sciences* 2002;6:37–46. [PubMed: 11849614]
- Zatorre RJ, Meyer E, Gjedde A, Evans AC. PET studies of phonetic processing of speech: Review, replication, and reanalysis. *Cerebral Cortex* 1996;6:21–30. [PubMed: 8670635]
- Zurif E, Swinney D, Prather P, Solomon J, Bushell C. An on-line analysis of syntactic processing in Broca’s and Wernicke’s aphasia. *Brain and Language* 1993;45:448–464. [PubMed: 8269334]



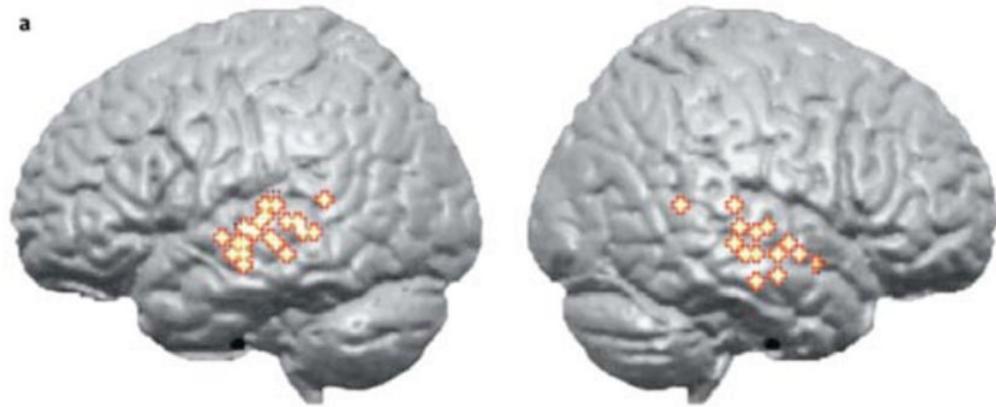
**Figure 1. Schematic of the functional anatomy of language processing**

Two broad processing streams are depicted, a ventral stream for speech comprehension that is largely bilaterally organized and which flows into the temporal lobe, and a dorsal stream for sensory-motor integration that is left dominant and which involves structures at the parietal-temporal junction and frontal lobe. ATL: anterior temporal lobe; Aud: auditory cortex (early processing stages); BA 45/44/6: Brodmann areas 45, 44, & 6; MTG/ITG: middle temporal gyrus, inferior temporal gyrus; PM, pre-motor, dorsal portion; SMG: supramarginal gyrus; Spt, Sylvian parietal temporal region (left only); STG: superior temporal gyrus; red line: Sylvian fissure; yellow line: superior temporal sulcus (STS). Adapted from (Hickok & Poeppel, 2007)



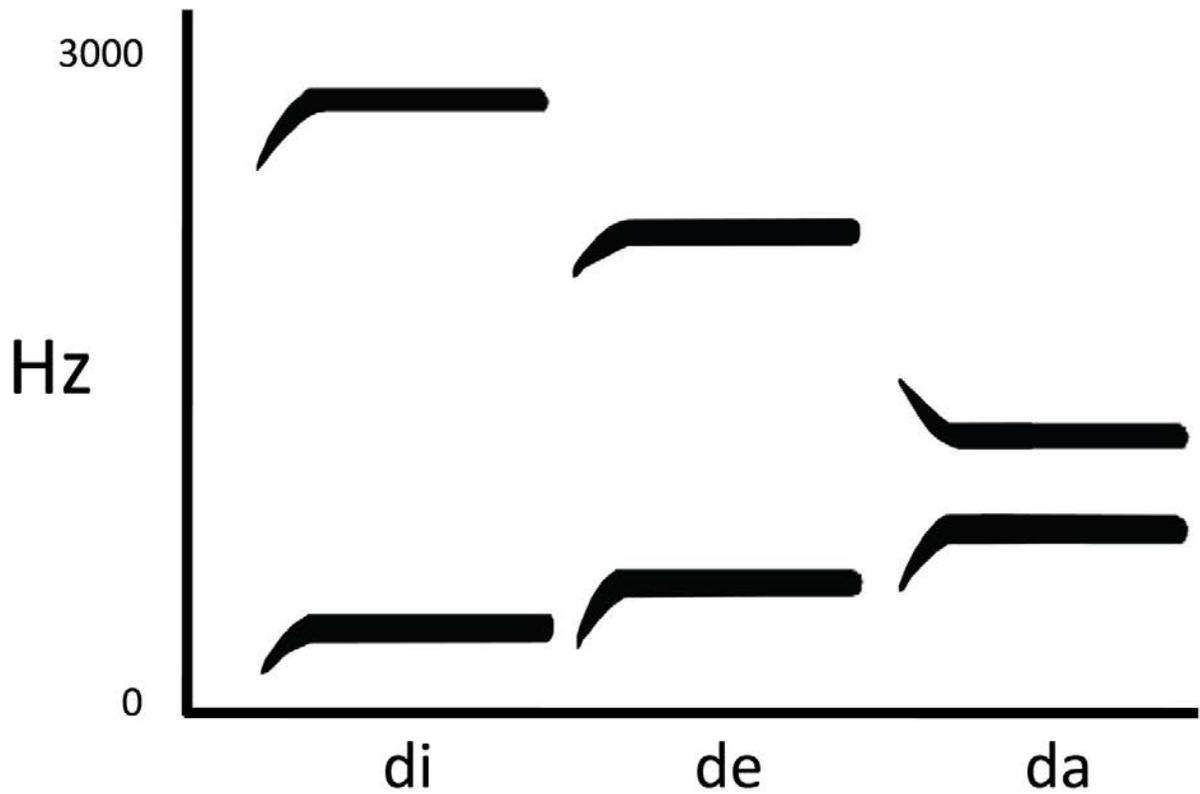
**Figure 2. Speech recognition in patients undergoing Wada procedures**

A sample stimulus card is presented along with average error rates of patients during left hemisphere anesthesia, right hemisphere anesthesia, or no anesthesia. Subjects were presented with a target word auditorily and asked to point to the matching picture. Note that overall performance is quite good and further that when patients make errors, they tend to be semantic in nature (selection of a semantically similar distractor picture) rather than a phonemic confusion (selection of a phonemically similar distractor picture). Adapted from (Hickok et al., 2008)



**Figure 3. Regions associated with phonological aspects of speech perception**

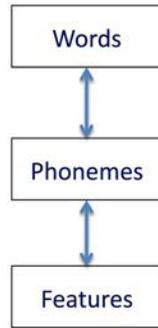
Colored dots indicate the distribution of cortical activity from seven studies of speech processing using sublexical stimuli contrasted with various acoustic control conditions. Note bilateral distribution centered of the superior temporal sulcus. Adapted from (Hickok & Poeppel, 2007).



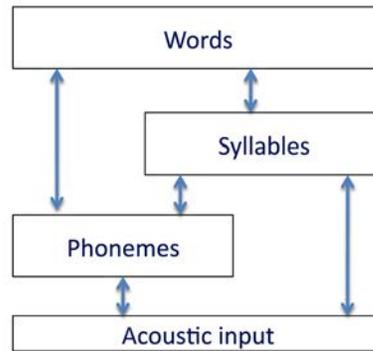
**Figure 4. Idealized spectrograms for three syllables**

The frequency sweep at the onset of each stimulus corresponds to the acoustic reflection of the /d/ sound in each syllable whereas the steady state portion corresponds to the vowels. Note that the acoustic variation associated with onsets despite the non-varying perception of the phoneme /d/. Adapted from (Liberman, 1957).

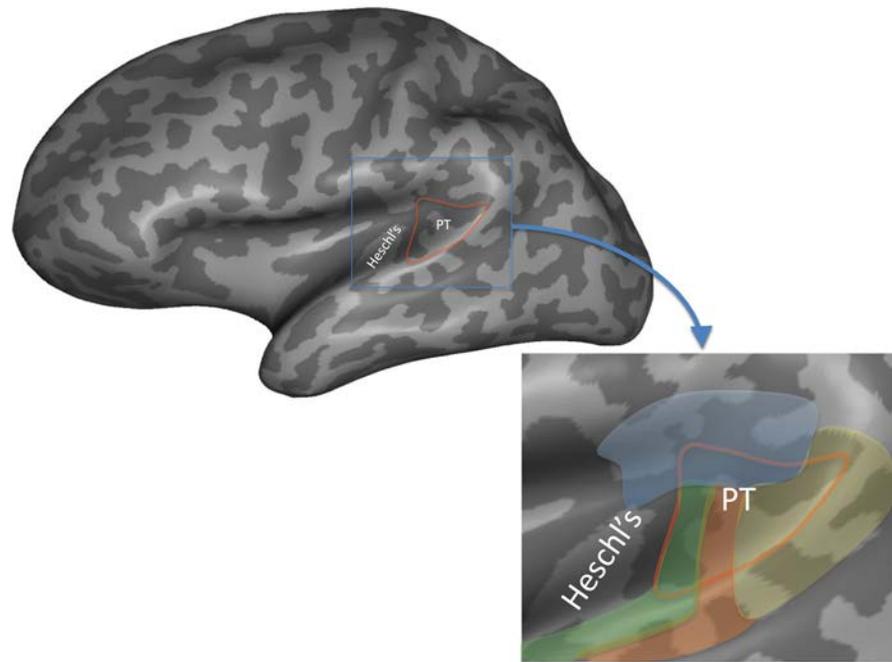
## A serial model



## A parallel model

**Figure 5. Serial versus parallel models of speech recognition**

The model on the left is typical of most models of speech recognition with some number of processing stages that are hierarchically organized and linked serially. The hypothetical model on the right depicts a possible architecture which allows for both hierarchical and parallel processing. Such an architecture, for example, allows speech to be processed at the level of the syllable, a unit that appears central in speech recognition (Luo & Poeppel, 2007) without necessarily accessing phoneme level units. Such an arrangement may avoid the lack of invariance problem (Sussman, 2002). The processing levels may be distributed across the two hemispheres in some fashion and may correspond to different temporal windows of integration (Boemio et al., 2005; Hickok & Poeppel, 2007).



**Figure 6. Location and cytoarchitectonic organization of the planum temporale**

The location of the planum temporale on the posterior supratemporal plane is indicated in red outline on an inflated representation of the brain which shows structures buried in sulci and fissures. The inset shows a close up of the planum temporale region. Colors indicate approximate location of different cytoarchitectonic fields as delineated by (Galaburda & Sanides, 1980). Note that there are four different fields within the planum temporale suggesting functional differentiation, and that these fields extend beyond the planum temporale. The area in yellow corresponds to cytoarchitectonic area Tpt which is not considered part of auditory cortex proper. Functional area Spt likely falls within cytoarchitectonic area Tpt, although this has never been directly demonstrated.

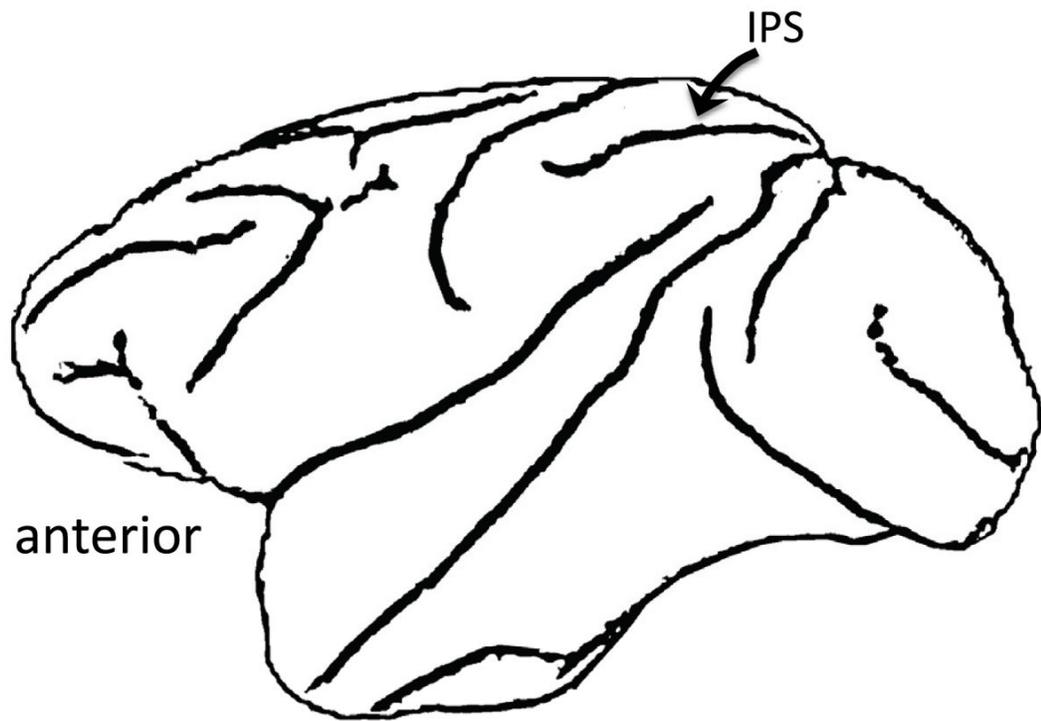
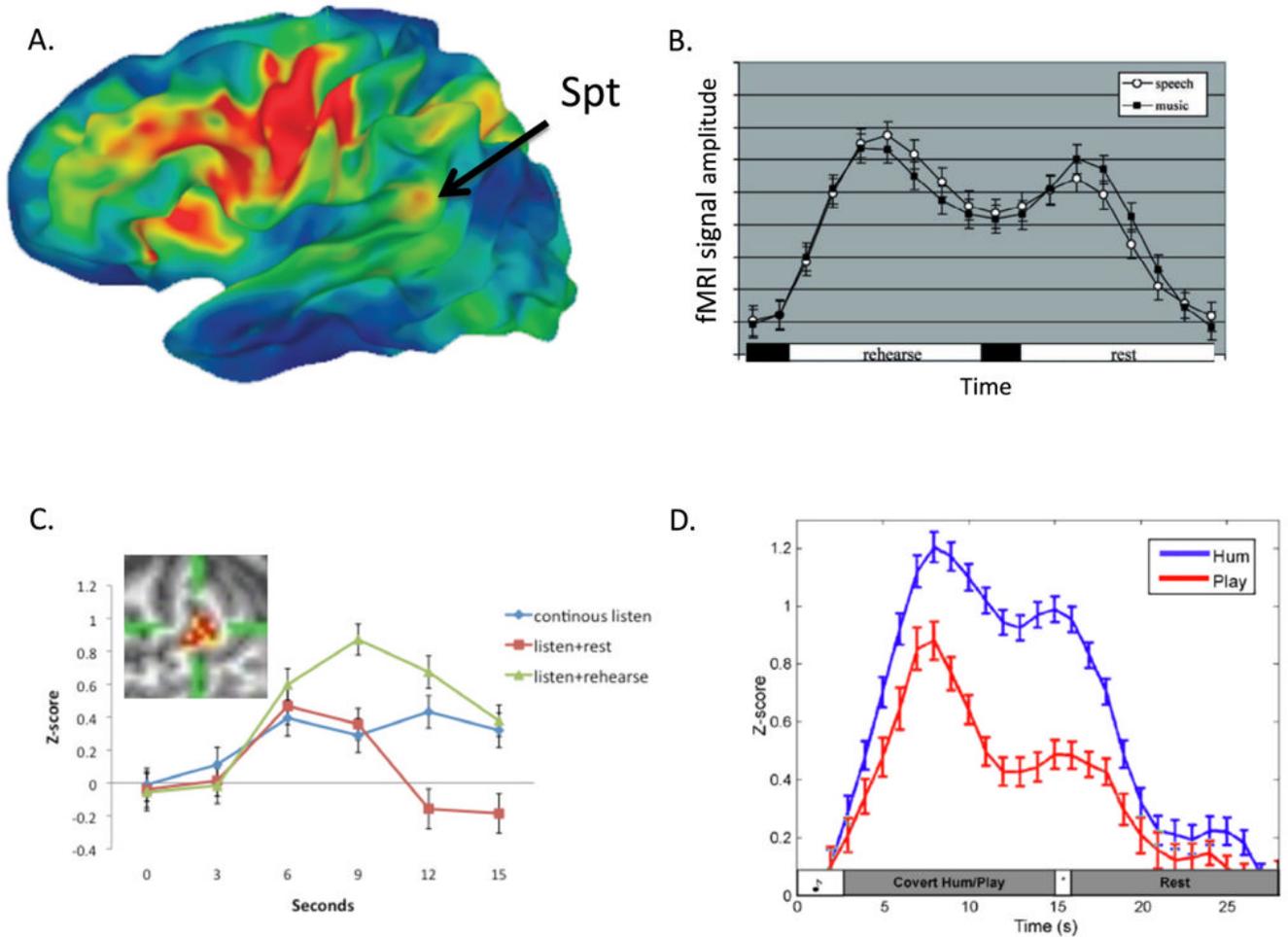


Figure 7. Diagram of macaque brain showing the location of the intraparietal sulcus (IPS)

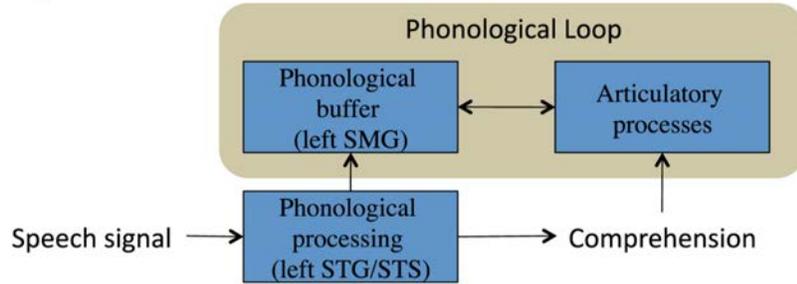


### Figure 8. Location and functional properties of area Spt

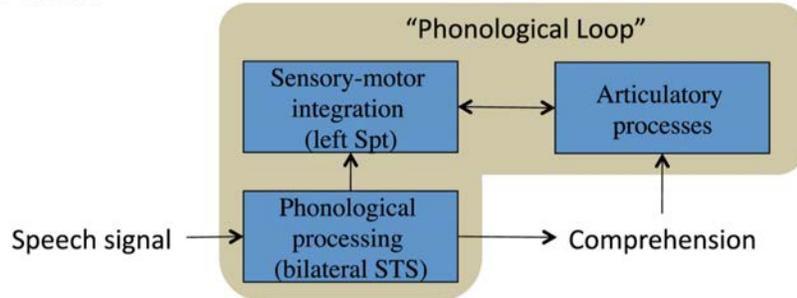
A. Probabilistic brain map of activated regions measured with fMRI in 102 subjects performing a sensory-motor speech task (listen to and then covertly rehearse speech). Colors indicate the percentage of subjects showing activation in a given region. Activity in Spt appears less robust than frontal regions because the region of activity is relatively small and because there is high anatomical variability in the posterior Sylvian region. The activation focus nonetheless provides a good indication of the typical location of Spt at the parietal-temporal boundary in the posterior-most aspect of the left Sylvian fissure. Adapted from Buchsbaum et al. 2009. B. Activation timecourse (fMRI signal amplitude) in Spt during a sensory-motor task for speech and music. A trial is composed of 3s of auditory stimulation followed by 15s covert rehearsal/humming of the heard stimulus followed by 3s seconds of auditory stimulation followed by 15s of rest. The two humps represent the sensory responses, the valley between the humps is the motor (covert rehearsal) response, and the baseline values at the onset and offset of the trial reflect resting activity levels. Note similar response to both speech and music. Adapted from (Hickok et al., 2003) C. Activation timecourse in Spt in three conditions, continuous speech (15s, blue curve), listen+rest (3s speech, 12s rest, red curve), and listen+covert rehearse (3s speech, 12s rehearse, green curve). The pattern of activity within Spt (inset) was found to be different for listening to speech compared to rehearsing speech assessed at the end of the continuous listen versus listen+rehearse conditions despite the lack of a significant signal amplitude difference at that time point. Adapted from (Hickok et al., 2009). D. Activation timecourse in Spt in skilled pianists performing a sensory-motor task involving

listening to novel melodies and then covertly humming them (blue curve) vs. listening to novel melodies and imagine playing them on a keyboard (red curve). This indicates that Spt is relatively selective for vocal tract actions.

### A. Phonological loop model

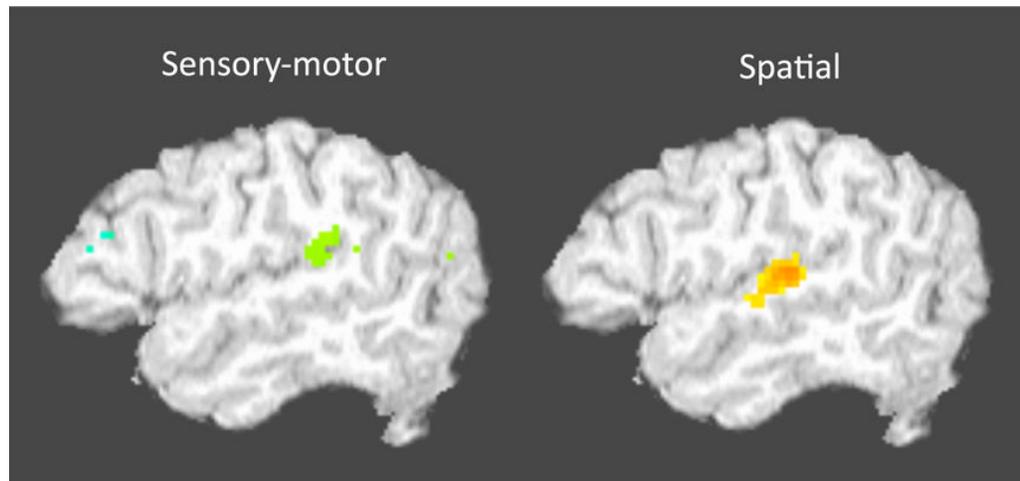


### B. Sensory-motor model



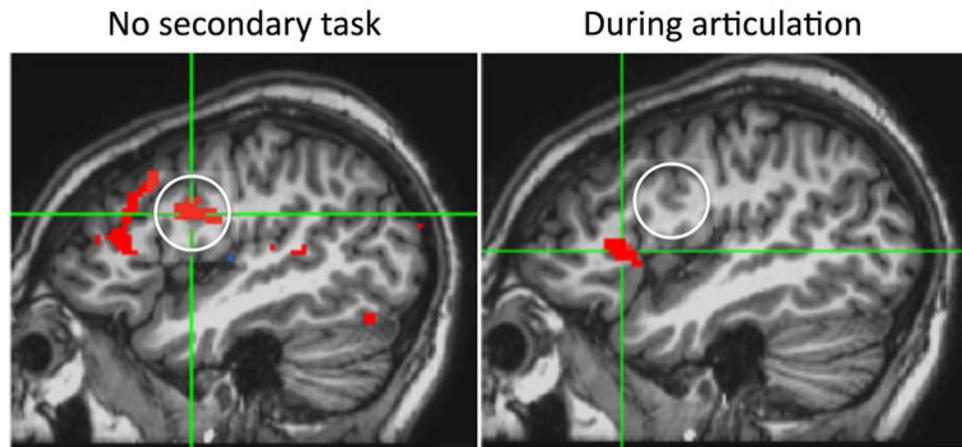
**Figure 9. Two models of phonological short-term memory**

A. Baddeley's phonological loop model (tan box) embedded in a simplified model of speech recognition. In this model, the phonological buffer (or "phonological store") is a dedicated storage device separate from phonological processing as it is applied in speech recognition. The contents of the buffer are passively stored for 2-3s but can be refreshed via articulatory rehearsal. B. A reinterpretation of the "phonological loop" within the context of a sensory-motor circuit supporting speech functions. Articulatory processes are the same across both models, but in the sensory-motor model the phonological "store" is coextensive with phonological processes involved in normal speech recognition and a sensory-motor translation system is interposed between the "storage"/phonological system and the articulatory system.

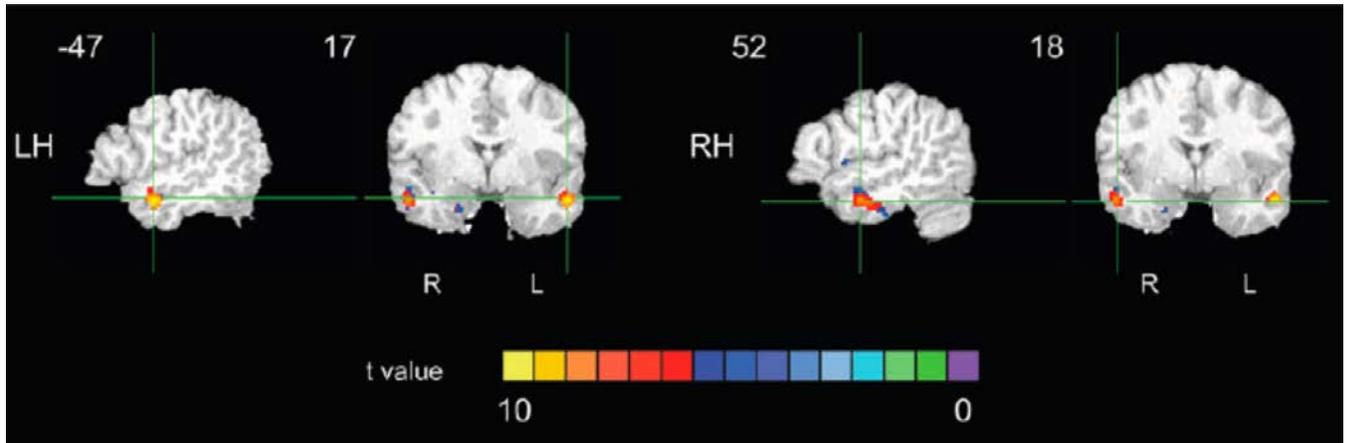


**Figure 10. Distinct activation foci for sensory-motor speech tasks and a spatial hearing task**  
Sensory-motor task is the standard listen and rehearse task and the spatial task is a moving sound source compared with a stationary sound source.

# Sentence Complexity Effect



**Figure 11. Effect of subvocal rehearsal on the sentence complexity effect in Broca's area**  
Left panel shows the standard sentence complexity effect (more activity for non-canonical word order sentences than for canonical word order sentences) in Broca's region. Right panel shows that most of this activity disappears once the effect of subvocal rehearsal is accounted for. Adapted from (Rogalsky, Matchinetal., 2008).



**Figure 12. The anterior temporal lobe “sentence area”**

Anterior temporal activity to auditorily presented sentences compared to lists of words.