Neuroscience Update

Where Are the Human Speech and Voice Regions, and Do Other Animals Have Anything Like Them?

The Neuroscientist
Volume 15 Number 5
October 2009 419-429

10.1177/1073858408326430
http://nro.sagepub.com

Christopher I. Petkov, Nikos K. Logothetis, and Jonas Obleser

---

Modern lesion and imaging work in humans has been clarifying which brain regions are involved in the processing of speech and language. Concurrently, some of this work has aimed to bridge the gap to the seemingly incompatible evidence for multiple brain-processing pathways that first accumulated in nonhuman primates. For instance, the idea of a posterior temporal-parietal "Wernicke's" territory, which is thought to be instrumental for speech comprehension, conflicts with this region of the brain belonging to a spatial "where" pathway. At the same time a posterior speech-comprehension region ignores the anterior temporal lobe and its "what" pathway for evaluating the complex features of sensory input. Recent language models confirm that the posterior or dorsal stream has an important role in human communication, by a reconceptualization of the "where" into a "how-to" pathway with a connection to the motor system for speech comprehension. Others have tried to directly implicate the "what" pathway for speech comprehension, relying on the growing evidence in humans for anterior-temporal involvement in speech and voice processing. Coming full circle, we find that the recent imaging of vocalization and voice preferring regions in nonhuman primates allows us to make direct links to the human imaging data involving the anterior-temporal regions. The authors describe how comparison of the structure and function of the vocal communication system of humans and other animals is clarifying evolutionary relationships and the extent to which different species can model human brain function.

Keywords: language; evolution; communication; primate; animal

---

> "[Paul Broca] was first to reject broad, indefinite expanses of the cortex as sites of speech areas and instead ventured to designate a very circumscribed, anatomically specific region as the seat of this function. As is well-known, he localized the faculty of speech to the posterior portion of the so-called third frontal gyrus."
>
> Carl Wernicke, 1874

In the 13 years between 1861 and 1874, the foundations of modern neuropsychology and cognitive neuroscience were established in no small part by Pierre Paul Broca, a French surgeon, and his contemporary Carl Wernicke, a Prussian psychiatrist (Broca 1861; Wernicke 1874). Many neuroscientists, even if they are not aware of the exact anatomical locations of the eponymous language regions, at least recognize these individuals for their enduring contributions.

Broca and Wernicke's observations provided strong arguments for both a localizationist approach that specific cognitive functions of the brain can be localized and a connectionist appreciation that functional modules are interconnected (Broca 1861; Wernicke 1874). The localizationist approach that they embraced was firmly tied to neuroanatomical knowledge of brain function. At the same time, Broca and Wernicke were skeptical of strict localization of brain function; yet, even Wernicke was quick to oversimplify Broca's stance (see the above quotation) and ignore his cautionary remarks that "the principle of localization by [gyral] convolution still rests on no certain basis" (Broca 1861). As a balance to strict localization, Wernicke, in particular, well reasoned that anatomical connectivity via neuronal tracts was instrumental for linking the speech-comprehension (Wernicke's) and speech-production (Broca's) centers of the brain (Fig. 1*a*, black arrow). He used this reasoning to predict new disorders involving impaired transcortical connectivity (Eggert 1977).

From the Max-Planck Institute for Biological Cybernetics, Tübingen, Germany (CIP, NKL); Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, UK (CIP); Imaging Science and Biomedical Engineering (ISBE), University of Manchester, Manchester, UK (NKL); and Max-Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany (JO).

The authors declare no competing financial interests.

We thank P. Belin and J. Fritz for useful discussions. Supported by grants from the Max-Planck Society and the Alexander von Humboldt Foundation.

Address correspondence to: Dr. Christopher I. Petkov, Institute of Neuroscience, Newcastle University, Framlington Place, Newcastle upon Tyne, NE2 4HH, UK (e-mail address: chris.petkov@ ncl.ac.uk).

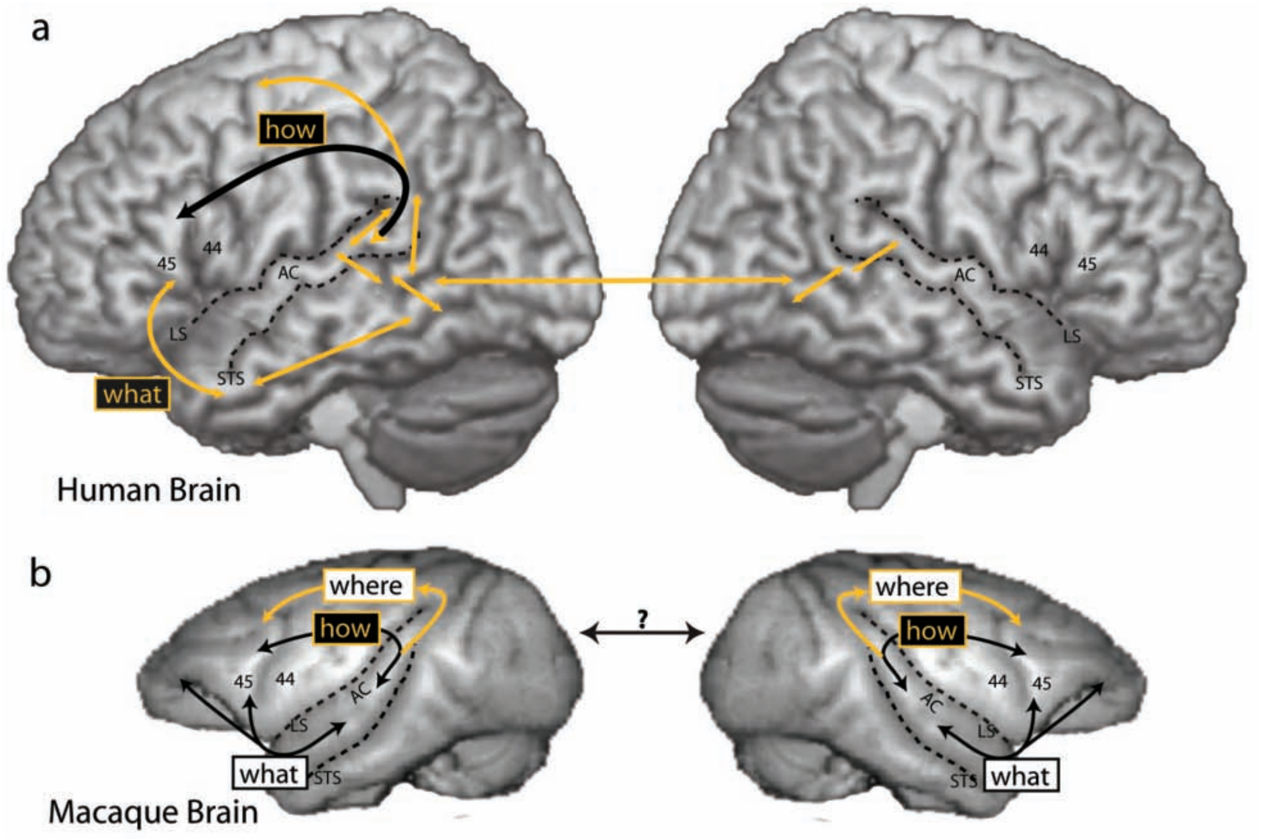

Figure 1. Original and updated pathways of vocal communication in humans and monkeys. The classical language pathway in the human left hemisphere (black arrow in a), which connects the posterior temporoparietal Wernicke's territory (speech comprehension) with the prefrontal Broca's territory (speech production). This model was incompatible with the dual "what" and "where" pathways in monkeys, shown in (**b**) with black lines and text. Using orange lines and text, we show the updated notion of the human pathways that support verbal language (**a**) and the pathways that process communication sounds in monkeys (**b**); see the text for details and the references that contributed to this illustration. Both models have adapted but differences are still apparent, some of the more obvious being that 1) the human brain is left lateralized whereas the monkey brain is assumed to be largely bilateral in its organization (but see Fig. 2 and Box 1), 2) it is unclear how many dorsal pathways there are in humans and monkeys, and 3) the "what" pathway in humans involves less of the regions in the superior and anterior parts of the temporal lobe. LS = lateral sulcus; STS = superior temporal sulcus; AC = auditory cortex; 44 and 45 = Brodmann areas.

The concepts outlined first by Broca and Wernicke still resonate in all neuroanatomical models of language, and ongoing work is clarifying the functional neuroanatomy of the network that supports human language, including the component systems dedicated to speech perception, comprehension, and production. However, variants on the classical language model needed updating and were, at first glance, incompatible with current ideas of multiple pathways of brain function that were developed in nonhuman primates (Hickok and Poeppel 2007; Scott and Johnsrude 2003). Efforts to come together on a neurophysiologically plausible model of language have required that the language-related and the pathways of processing models adapt, with new hypotheses for testing becoming clearer in the process (see Fig. 1; Hickok and Poeppel 2007; Scott 2005; Scott and Johnsrude 2003). We think that these efforts can now come full circle and return to nonhuman animals to help us bridge the gap between the neurobiology of human and animal communication. To guide such efforts we review exemplary issues in the human and animal literature on the neuronal bases of communication and, in particular, attempt to reconcile the recent nonhuman primate imaging work that used communication sounds to compare their results with those obtained from imaging humans.

As we will illustrate, at the moment successful comparative endeavors to study the neuronal basis of language perception and production are mostly restricted to the signal- and stimulus-bound aspects of language such as voice and speech processing, the neurobiological capacities termed by Hauser and others (2002) as belonging to the "language faculty in the broad sense." Although it is generally accepted that the more abstract linguistic skills, such as complex syntax, may be unique to humans, their simpler forms might also be found in other animal species (Fitch and Hauser 2004; Hauser and others 2002). In any case, the grammatical aspects of our language, for instance, depend upon stimulus-bound processing (Wernicke 1874), which some believe originate from simpler capacities that have been conserved in some nonhuman animal species (e.g., Fitch and Hauser 2004; Friederici and others 2006; Hauser and others 2002).

The Human Language Regions: Revolution followed by Constant Reform

Modern lesion and brain imaging work (e.g., Damasio and Geschwind 1984; Petersen and others 1988) sets the stage for questioning the extent of the regions supporting speech and language processing, the number of pathways involved, and the traditional interpretations of Broca and Wernicke's seminal works. Accumulating knowledge makes it clear that speech and language processing involves several territories composed of cortical and subcortical regions and relies on multiple processing pathways (Dronkers 1996; Hickok and Poeppel 2007; Liebenthal and others 2005; Naeser and others 1982; Obleser and others 2006; Rimol and others 2005; Scott and others 2000; Wise and others 2001).

For one, considerable knowledge has been acquired about the area that is commonly referred to as Broca's region: how the inferior frontal cortex is delineated probabilistically and cytoarchitectonically (e.g., Amunts and others 1999) and to what degree it should be extended to comprise a language subsystem by including other cortical and subcortical structures. Also, on the basis of Broca's interpretations and of new fMRI evidence (e.g., Dronkers 1996; Dronkers and others 2007), one can now ask how Broca's region interacts with other structures to contribute toward language comprehension (e.g., Friederici and Kotz 2003; Thompson-Schill and others 1997), and whether monkeys have neuroanatomically (e.g., Petrides and Pandya 2002) or functionally (Rizzolatti and Arbib 1998) any similar subsystem.

Moreover, the other classical language region, Wernicke's region, is also conceptually adrift albeit it drew less neuroscientific attention than Broca's region. This might be so because Wernicke, although oversimplifying Broca's conclusions on the location of the "motor speech" center, as he called it, was himself less specific about where a speech-comprehension ("sensory speech") center would reside. Thus Wernicke's region was either too focally "localized" by other investigators to the temporoparietal junction, involving the angular gyrus and posterior superior-temporal cortex, or it was considered too broad to be meaningful (Wise and others 2001), although perhaps better in line with Wernicke's writings.

Lesions to Wernicke's region are thought to result in speech comprehension disorders, together called comprehension aphasia. Yet, it is of note, and speaks to the persisting difficulties with pinpointing the language-related regions, that tremendous variance occurs in the actual lesion sites and patients often present different symptoms that together encompass the syndromes of comprehension or production aphasia (e,g., Bates and others 2003).

Recent studies in humans have observed that much of the anterior as well as the posterior stretches of the superior and middle temporal lobe are involved in a variety of speech-related processes (see Fig. 2*a* for sublexical processing, and for lexical processing, see Scott and others 2000; Scott and Johnsrude 2003; Wise and others 2001). This is consistent with the idea that speech processing cannot be strictly localized to a posterior temporal region or pathway. It is also very much in line with Wernicke's interpretations of the lesion sites in aphasic individuals. Wernicke himself contemplated the possibility that the sensory-speech center (perhaps better termed a "subsystem") involved many of the cortical and subcortical regions around the Sylvian fissure (lateral sulcus), without specifying their extent and location, for example, whether they were more anterior or posterior in the temporal lobe (Wernicke 1874).

Additionally, although most language models assume lateralization, few studies have explicitly tested and showed lateralization effects (see Fig. 2 and Table 1). Wernicke himself concentrated almost exclusively on the left hemisphere; nonetheless he was careful to emphasize that many comprehension aphasia patients did have bilateral lesions and that the language lateralization to the left hemisphere was often a gross generalization from right-handed cases (Eggert 1977; Wernicke 1874). Current thinking is that at least the initial stages of speech processing are bilaterally distributed, with lateralization becoming more prominent at later processing stages (Hickok and Poeppel 2007).

Although the specific linguistic and nonlinguistic functions of the various language-related regions of the brain and their connectivity remain under active study, it is notable that the basic concepts outlined by Broca and Wernicke still guide those endeavors. As we will consider next, what is becoming clearer is that Broca's and Wernicke's areas have been redefined as territories of interconnected brain areas, heavily drawing upon and now helping to guide the concepts of the multiple processing pathways that were derived from animal studies.

Multiple Networks and Pathways of Processing Speech in Humans

Rather than the classical two human speech areas, the evidence for more extensive cortical regions involved in different aspects of speech perception and comprehension that has accumulated better relates to the idea of language processes relying on multiple networks or parallel processing pathways in the brain (Hickok and Poeppel 2007; Scott and Johnsrude 2003). Originally, visual studies in monkeys suggested that the brain processes different aspects of sensory stimuli largely in parallel, with a

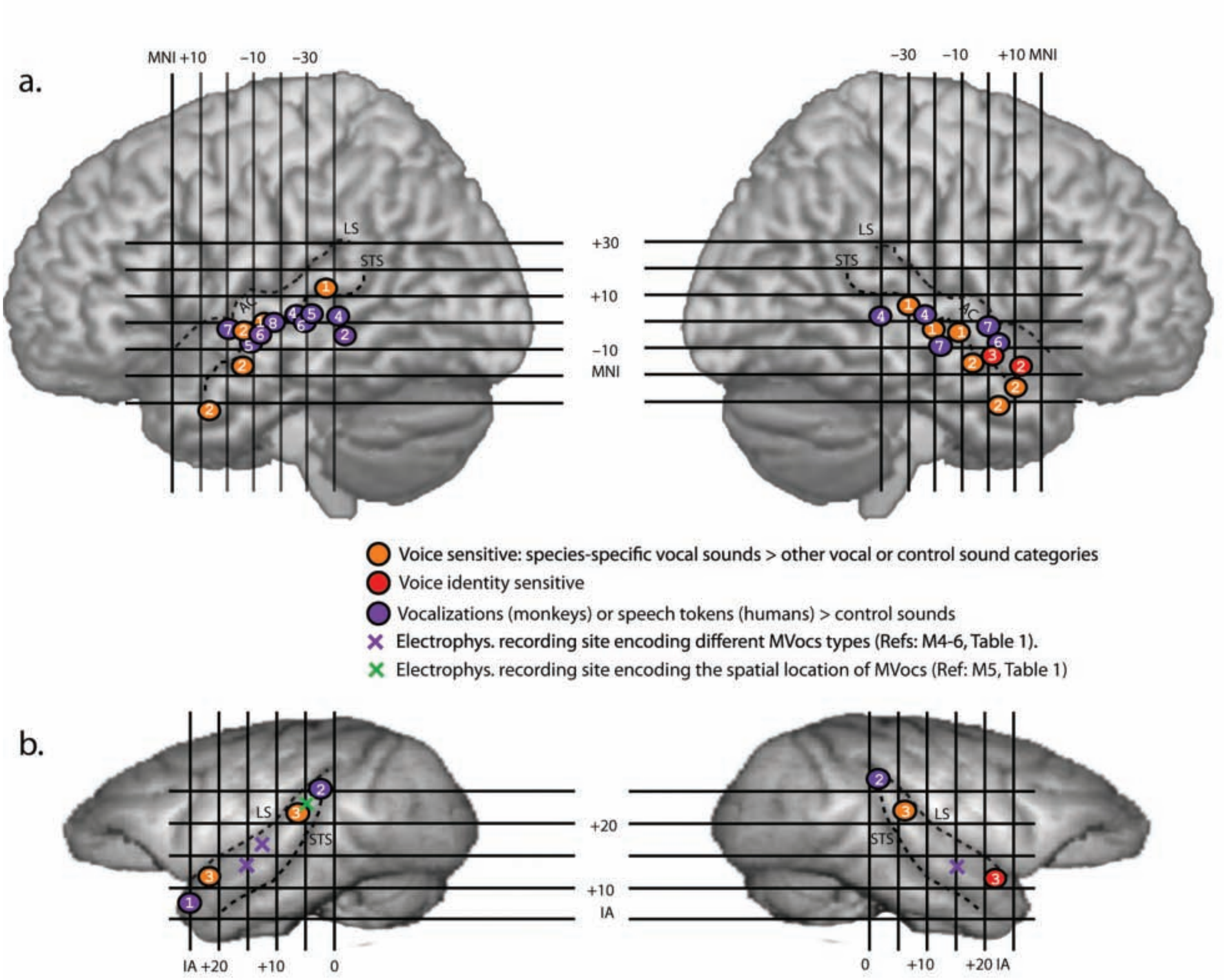

Figure 2. Comparative summary of human and macaque processing of species-specific communication sounds in the superior temporal lobe. (**a**) Summarized results from several human sublexical and voice processing imaging studies shown as labeled dots. These can be compared with the results from macaque studies that have used the vocalizations of conspecifics (**b**). Here we focus on the results obtained in the superior temporal lobe outside of primary auditory cortex. See Table 1 for a detailed description including the references for each study that was used. Note how at least the peak of activity selective for the processing of voice content in monkey vocalizations appears to be more superior on the temporal lobe in (**b**) than the processing of human voice content in the human brain (**a**); compare the anatomical location of the red circles and see the text for details. MVocs = macaque-monkey vocalizations; MNI = Montreal Neurological Institute coordinates; IA = interaural coordinates (see Table 1).

ventral "what" pathway dedicated to analyzing the features of visual stimuli and a dorsal pathway involved in the analysis of "where" visual objects are located in or how they translate through space (Ungerleider and Mishkin 1982). This idea has since found support in the central auditory system (Fig. 1), first in studies with monkeys (Rauschecker 1998; Romanski and others 1999) and then with humans (Alain and others 2001) and other mammals (Lomber and Malhotra 2008).

Given the complex sensory features of vocal communication signals, a parsimonious theory would consider that any neural circuitry of speech perception and comprehension needs to engage the ventral or anterior "what'" pathway. For instance, although the analysis of acoustic "objects" encompasses all sorts of environmental sounds, some of the analysis conducted in the "what" pathway would be instrumental in speech perception and language comprehension because it disambiguates speech tokens and details the features of speech sounds (Griffiths and others 2004). Recent evidence has been accumulating that the human anterior temporal lobe is involved in several speech processing functions (see Fig. 2*a*; Hickok and Poeppel 2007; Liebenthal and others 2005; Obleser and others 2006; Rimol and others 2005; Scott and others 2000; Wise and others 2001). This anterior temporal processing of human speech sounds neuroanatomically seems to correspond to the processing of the features in monkey vocalizations in the anterior parts of the monkey auditory cortex (see Fig. 2*b*; Rauschecker 1998; Tian and others 2001). As for a

Table 1. Study Details

Primate	Method	Fig. 2 Label	Task	Comparison	Coordinates (x,y,z) Brain LH	RH	Region
Humans (H)	fMRI	1.	Passive stim.	HVoc > Nonvocal	−63, −14, −1	64, −13, −2 53, −20, −2	Mid. STS
					−40, −39, 12	57, −31, 5	Post. STG
		2.	Voice recogn.	Voice ident. > Speech	[?]	54, 13, −17	Ant. STS
				Speech > Voice ident.	−58, −46, −6	[?]	Post. MTG
				Voice > Control	−64, −15, −4	58, 10, −24 58, 4, −32	Ant. STS/MTG
					−61, −8, −18	−64, −6, −15	Ant. Mid. STS/MTG
					−55, 8, 32		Ant./Mid. STG
		3.	Passive stim.	Voice identity	[?]	59, 2, −9	Ant. STS
		4.	Discrimination	Phonemes > Nonphonemes	−61, −25, 3	61, −25, 3[a]	Post. STG/STS
					−57, −41, −2	57, −41, −2[a]	Post STS/MTG
		5.	Discrimination	Phonemes > Inverted spectrum	−61, −8, −4	[√]	Ant./Mid. STS
					−57, −32, 2	[√]	Post. STS
		6.	Detect sound repetition	Syllables > Noise	−61, −8, −4	59, 4, −10	Ant./Mid. STS
					−57, −32, 2	[?]	Post. STS
		7.	Vowel detection	Vowels > Band-passed noise	−64, −2, −2	62, 0, −4 62, −18, −11	Ant. STG
		8.	Intelligible, not intelligible	Consonants > Inverted spectrum	−78, −18, 0	[?]	Mid. STS
Macaques (M)	PET	1.	Passive stim. (Scan anesth.)	MVoc > controls	−18, 25, 8	[√]	Temporal pole
		2.	Passive stim. (Scan awake)	MVoc > controls	−27, 2, 25	27, 2, 25	Tpt
	fMRI	3.	Passive stim. (Scan awake or anesth.)	MVoc > controls	−20, 6, 22	20, 6, 22	Primary or belt AC
					−21, 21, 12	21, 21, 12	Anterior STP
				Voice ident.	[?]	21, 21, 12	Anterior STP
	Electrophys.	4.	Anesth.	MVoc preference	−27, 13, 17	[?]	Anterior belt
		5.	Anesth.	Tuned for MVoc	−27, 13, 17	[?]	Anterior belt
				Spatial tuning	−25, 5, 24	[?]	Caudal belt
		6.	Passive stim. (Awake)	MVoc coding	−28, 15, 14	28, 15, 14	STG

References for the human studies: 1. Belin P, et al. 2000. Nature 403:309–12. 2. von Kriegstein K, et al. 2003. CBR 17:48–55. 3. Belin P, Zatorre RJ. 2003. Neuroreport 14:2105–9. 4. Dehaene-Lambertz G, et al. 2005. Neuroimage 24:21–33. 5. Liebenthal E, et al. 2005. Cereb Cortex 15:1621–31. 6. Rimol LM, et al. 2005. Neuroimage 26:1059–67. 7. Obleser J, et al. 2006. Hum Brain Mapp 27:562–71. 8. Obleser J, et al. 2007. Cereb Cortex 17:2251–7.

References for the macaque studies: 1. Poremba A, et al. 2004. Nature 427:448–51. 2. Gil-da-Costa R, et al. 2006. Nat Neurosci 9:1064–70. 3. Petkov CI, et al. 2008. Nat Neurosci 11:367–74. 4. Rauschecker JP, et al. 1995. Science 268:111–4. 5. Tian B, et al. 2001. Science 292:290–3. 6. Russ BE, et al. 2008. J Neurophysiol 99:87–95.

Details on the studies that are summarized in Fig. 2. The human coordinates are the reported peak of the activity cluster, as reported in or converted to the coordinates of the Montreal Neurological Institute (MNI) standard. The monkey coordinates use the Frankfurt zero system standard (interaural line of zero and the orbitomeatal plane), but are approximations of the reported peaks of activity or the centers of the recording sites because the original studies did not report coordinates.

[?] = Lateralization effect was not explicitly tested or site recorded. Possible subthreshold contribution of this hemisphere; [√] = Lateralization effect was explicitly tested and this hemisphere does *not* significantly contribute to this particular task and comparison, relative to the contralateral hemisphere. HVoc = human vocal sounds; MVoc = monkey vocalizations; STS = superior-temporal sulcus; STG = superior-temporal gyrus; MTG = middle-temporal gyrus; STP = superior-temporal plane; Ant. = anterior; Post. = posterior; anesth. = anesthetized; stim. = stimulation; recogn. = recognition; ident. = identity.

[a] These RH regions are subthreshold in the original report, but we have added them here because the authors found a nonsignificant lateralization effect for the reported comparison in these regions (see human reference 4).

posterior or dorsal pathway in speech comprehension (involving the caudal temporoparietal regions and the arcuate and superior longitudinal fasciculi for connectivity with the frontal cortex), it has been plausibly suggested that the "where" pathway could serve a "how-to" role in motor preparation (Hickok and Poeppel 2007; Warren and others 2005). An alternative to the dual pathway models (what/where or what/how) is the notion that several pathways exist, such as those needed to evaluate the "how," "what," and "where" aspects of the sensory input (see Fig. 1; Scott 2005; Scott and Johnsrude 2003).

However, a close correspondence between the dorsal pathway(s) in humans and monkeys, linking temporal to prefrontal territories, is uncertain: There are considerable differences in brain connectivity when comparing human and nonhuman primate brains.

Recently, Rilling and colleagues (2008) conducted a comparative human, chimpanzee, and monkey study using MR-based diffusion-tensor imaging to evaluate neuronal connectivity in all three species. The authors evaluated the neuronal tracts linking the prefrontal (Brodmann areas 44, 45, part of "Broca's territory" in humans) and the superior and middle temporal cortex (part of "Wernicke's territory" in humans). They observed that the connectivity between these regions for macaque monkeys depended mainly on a more ventral, extreme capsule pathway rather than on the dorsal, arcuate, and superior longitudinal fasciculi; the neuronal tracts, which are more evident in humans. Interestingly, chimpanzees show components of both the human and macaque connectivity, suggesting that the human tracts that are presumably key for connecting the language-related regions evolved gradually within the primate order (Rilling and others 2008).

Schmahmann and colleagues (2007) independently noted, using a related "diffusion-spectrum imaging" technique combined with anatomical tractography, that the dorsal monkey "arcuate fasciculus," if it exists, is unlikely to project to the same Brodmann areas (BA 44 and 45). The authors noted that the tract or tracts most likely linking temporal and prefrontal cortex in monkeys are the tracts of the extreme capsule (in agreement with Rilling and others 2008) and/or the middle longitudinal fasciculus (MiLF; Schmahmann and others 2007).

At first sight, the structural homologies involving the dorsal networks appear limited. However, the language-related regions and their anatomical connectivity are under active pursuit in humans (Friederici and others 2006). Additionally, further comparative data involving both structure and function will be needed to support the claim of limited homologies in the dorsal pathways rather than the alternative hypotheses. It is an open question how many distinct regions in the brains of primates form neuronal networks that are evolutionarily related to those of humans that we currently identify, and especially which are the precursory networks upon which human language evolved. Yet, the notion that the human brain underwent considerable differentiation to support language makes it likely that a number of evolutionarily conserved regions have differentiated and repositioned in humans to occupy different anatomical locations. Once repositioned, some of these regions in humans may have had to—or merely appear to—rely on different neuronal tracts for their connectivity, in relation to the ones present in the extant primates. We consider the possibility of conserved function but different anatomical location in the next section with an example from a recent macaque fMRI study and a summary of several human and monkey studies (Fig. 2 and Table 1). Our comparisons illustrate a certain level of correspondence between human and macaque brains for the processing of species-specific vocal sounds for communication, especially regarding the ventral or anterior-temporal pathway.

Precursors to the Language Regions in Nonhuman Animals

Humans have unique verbal and nonverbal abilities for communication. But there are behavioral commonalities in the use and production of communication signals across many vertebrate species, suggesting that components of the human communication system would be evident in other species (Hauser and others 2002). To unravel the neurophysiological basis for vocal communication, for instance, requires the use of species-specific vocalizations within a neuroethological context, which considers both how animals use communication signals and how their brains might have adapted to help them to survive and meet their social needs (Ghazanfar and Hauser 2001; Griffiths and others 2004; Suga 1992).

In birds and bats, for instance, there is clear specialization of the auditory cortex for the processing of species-specific communication signals. The bat auditory cortex is specialized for an ability that humans lack, namely echolocation. Yet, these studies reveal neuronal mechanisms that sculpt the responses to species-specific vocalizations, which are thought to be broadly conserved (e.g., Suga and others 1979). Singing birds seem like a good model for human vocal learning inasmuch as there is now a large literature describing how avian brains rely on posterior sensory centers that are necessary for song learning and comprehension and anterior premotor centers necessary for song production (for a review, see Jarvis 2004). Accordingly, singing birds seem to exhibit the closest animal variants to Wernicke and Broca's territories that have been described. Given that avian and mammalian brains differ considerably (but see Jarvis and others 2005), these are likely independent adaptations to support vocal learning and production within a broadly conserved posterior to anterior sensory-motor axis of brain organization.

Nonhuman primates have the closest evolutionary relationship to humans. Thus, many have been looking to monkeys and apes for homologies in brain function that would reveal the origins of human language.

A number of electrophysiological studies in monkeys have used species-specific vocalizations as meaningful sound stimuli and reported strong responses to these sounds in primary auditory cortex, surrounding belt regions, parietal cortex, and prefrontal cortex. As examples, in monkey temporal and prefrontal regions there are strong neuronal responses to species-specific vocalizations (e.g., Funkenstein and Winter 1973; Rauschecker and others 1995; Tian and others 2001; Wang and Kadia 2001), neuronal responses can encode or decode different call types (Romanski and others 2005; Russ and others 2008) and some regions in the prefrontal cortex could serve an important role in evaluating the meaning of conspecific vocalizations (Gifford and others 2005).

Recent imaging studies of monkeys have aimed to complement the electrophysiological approach with a balanced and more global evaluation of the many regions of the brain. Also, by using the same technique, the monkey imaging can be directly compared with imaging data obtained in humans. Unfortunately, the first few imaging studies with PET and fMRI in macaques provided conflicting results, which we attempt to reconcile in our effort to relate them to the processing of vocal and speech sounds in humans.

Despite the recent developments in human speech and language research that point beyond the classical Broca-Wernicke areal distinctions, a recent macaque PET imaging study claimed that monkey homologs of the classical human language areas were revealed: A caudal temporal area and a ventral prefrontal area, presumed to be the homologs of Wernicke's and Broca's areas, respectively, responded strongly to macaque vocalizations (Gil-da-Costa and others 2006). However, the initial enthusiasm was tempered because in a previous report from the same study the authors noted that the monkey vocalizations were effectively driving many monkey-brain regions and networks (Gil-da-Costa and others 2004), as might have been expected on the basis of the brainwide responsiveness to complex sounds (Poremba and others 2003). Further, the stimulation repertoire of the above study consisted of a limited number of monkey vocalizations (only "coos" and "screams") and—spectrotemporally—simple (nonbiological) sounds for comparison. There was also little consistency with the observations from an earlier macaque PET imaging study, which reported that an anterior, rather than a posterior, temporal-lobe region unilaterally represents the monkey vocalizations, with stronger responses to these species-specific vocalizations in the left hemisphere (Poremba and others 2004).

Given that even the human Wernicke's and Broca's regions have turned out to be not so strictly localized to a single caudal temporal or ventral prefrontal area, more convincing evidence for precursors to the human language regions might come from the framework of multiple processing pathways, and from considering how the features of species-specific vocalizations are analyzed along these paths.

A recent high-resolution fMRI study in macaques (Petkov and others 2008) can largely reconcile the previous macaque PET imaging results. In this case, the authors used a design that was similar to that of Poremba and others (2004) where they compared the activity response to a large set of macaque vocalizations over other spectrotemporally complex sound categories, including other animal vocalizations. The fMRI results showed a strong preference for the acoustics of species-specific vocalizations along both anterior and posterior superior-temporal regions, bridging the observed anterior- (Poremba and others 2004) and posterior-temporal (Gil-da-Costa and others 2006) results with PET (Fig. 2*b*). However, unlike the "Wernicke-like" area emphasized by Gil-da-Costa and colleagues, the more posterior activations preferring macaque vocalizations were—with high-resolution fMRI—localized closer to the primary auditory cortex (Fig. 2*b*). This makes it difficult to directly relate the results of these two PET and fMRI studies, but one interesting—although highly speculative—possibility is that the PET study better showed higher level regions, that presumably belong to the dorsal "how" pathway, because the animals were free to vocalize during the acoustical stimulation (Gil-da-Costa and others 2006). In the fMRI study, in contrast, different result were obtained perhaps because the animals were not able to vocalize; the authors used sensors to instruct the animals not to move their jaw and body during each scanning trial because body movements disrupt the quality of the imaging data at high-magnetic fields (Keliris and others 2007; Petkov and others 2006; Petkov and others 2008). Further monkey imaging with more sophisticated experimental designs for controlling the animals' behavior is needed to evaluate the contribution of the dorsal pathways in communication sound processing because we expect both monkey and human study results to depend critically upon the precise involvement of vocal or speech comprehension and production (cf. Hickok and Poeppel 2007; Wilson and others 2004).

In the fMRI study of Petkov and others (2008), the more reliable anterior temporal lobe regions that prefer species-specific vocalizations belong to the "what" processing pathway and show a close functional correspondence to a now well-known human brain region, also in the anterior temporal lobe (Belin and Zatorre 2003; Belin and others 2000; von Kriegstein and others 2003). The human and monkey regions in the anterior temporal lobe are termed "voice" regions because they

Box 1. Pathways Ahead: Comparative Studies of the Communication Networks

For guiding future endeavors we articulate several general questions or hypotheses for which the comparative method could be used to clarify the relationship between the neural systems that support human language and those that support the communication abilities of other animals.

1. Multiple pathways of processing hypothesis. Vocal/verbal communication signals are processed in multiple largely homologous pathways (at least the "what," "where," and "how-to" pathways) although some of the brain regions in humans have differentiated considerably and may thus rely on different or seemingly different (repositioned) neuronal tracts for connectivity.
2. Shared voice and speech/vocalization processing. Are the neural substrates for voice and speech processing shared in humans? Figure 2*a* suggests that this is the case by the proximity of the human voice and speech processing. Also, what are the analogues of speech-like processing in nonhuman animals?
3. Lateralization of function. The initial sensory processing of communication signals occurs bilaterally and becomes more left or right lateralized as the computational demands on the processing increase (computational demands hypothesis, e.g., see Hickok and Poeppel 2007) or as different functions become segregated (modularity of function hypothesis). Comparative studies will need to directly test left vs. right lateralization hypotheses, because lateralization claims are often improperly substantiated (see Fig. 2 and Table 1).
4. Early cross-sensory influences hypothesis. Human language need not be verbal and some have questioned whether neocortex is unisensory at all (Ghazanfar and Schroeder 2006). Comparative evaluation of the cross-sensory influences along the different processing pathways will clarify whether additional functional homologies can be described or if there are striking differences across the species in how brains process cross-sensory signals.

are sensitive to both the vocal identity (voice) of different conspecific individuals and to the vocal signature of species-specific vocalizations, the so-called voice of the species (see Fig. 2, Table 1; Belin and Zatorre 2003; Belin and others 2000; Petkov and others 2008; von Kriegstein and others 2003).

Although the interaction of human voice processing with speech and language processing is still a matter open for debate,[1] the different systems are all constituent aspects of communication. The human and monkey imaging data on the voice regions reveal a reasonable high-level homology in the processing of species-specific vocal features and thus link the human and macaque vocal processing that occurs in the anterior temporal regions of these primates. This is illustrated in Figure 2 and Table 1, where we also summarize the results obtained on the superior temporal lobe from several fMRI studies of human sublexical and voice processing (Fig. 2*a*). We compare these human results to macaque imaging and electrophysiology studies that have used communication sounds, focusing again on the superior temporal lobe (Fig. 2*b*). This synoptic view reveals that the processing of monkey and human vocalizations clearly involves both anterior and posterior superior-temporal regions or multiple processing pathways.

Our comparison in Figure 2, however, also shows an interesting difference in the sites that process human or monkey vocal and communication sounds, which to our knowledge, was first noted in the monkey fMRI study by Petkov and colleagues (2008). Namely, the human sites of the peaks in activity are lower in the temporal lobe (Fig. 2). This goes along with the observation that although the monkey brain seems to process monkey voices just as the human brain processes human voices, the anatomical locations of the human and monkey "voice" regions seem to differ (Petkov and others 2008). A rough comparison of the location of the monkey voice region, which is on the superior temporal plane (Petkov and others 2008), with the location of the reported human voice region (Belin and Zatorre 2003; von Kriegstein and others 2003), reveals that the human region is located lower in the temporal lobe, within the upper bank of the superior-temporal sulcus (see Fig. 2 and Table 1). Although it is unclear whether this is specific to basic voice processing and how these regions may have repositioned, the observation could have implications for understanding the evolutionary basis of the human speech-processing regions.

For instance, if additional comparative data support the observation of similar function but different anatomical position or connectivity, the data could clarify how the neighboring speech-processing regions in humans were affected by the evolutionary differentiation of the temporal lobe. Comparative study of human and macaque brain anatomy already shows some evidence for a global differentiation of the human superior temporal lobe in comparison to the macaque brain (Fullerton and Pandya 2007) rather than for some parts of the temporal lobe to have differentiated more than others. Thus, a careful comparative study of the anatomical positions of brain regions with conserved function, such as the voice region, would likely also reveal how the nearby speech-processing regions in humans also differentiated away from the superior temporal plane (as Fig. 2*a* suggests), even if many of these have no direct counterparts in extant species. However, if further largely homologous brain regions are revealed in other primates, it is reasonable to expect that, given the considerable differentiation of the human brain, some of these regions in humans now rely on different or repositioned neuronal tracts for their connectivity by comparison to nonhuman species.

Coming Together on the Path Ahead

The careful work and far-reaching insights of Broca and Wernicke gave us some of the first fascinating examples of the regional functional specialization in the brain. Their results and conclusions, even if often overinterpreted, paved the way for elucidating the neural basis of human communication. Interestingly yet, these investigators were much less specific about the site and extent of the human sensory-speech and motor-speech centers than their successors. It has only recently become clear that Broca's and Wernicke's areas are not necessarily topologically compact, but rather they form subsystems consisting of a number of interconnected functional regions around the lateral sulcus, including subcortical regions (Fig. 1). Studies carefully controlling the speech and linguistic properties of stimuli while considering the behavioral task are clarifying the precise function and involvement of the many human communication-related regions (e.g., Friederici and others 2006; Hickok and Poeppel 2007; Marslen-Wilson and Tyler 2007; Scott and others 2000). Some of this work has also been able to better relate to the ideas of parallel processing that were developed in animal models.

Just as in other research fields, understanding how the brain enables linguistic communication requires an integrative approach that takes into account homologies between humans and other species, as well as using combinations of noninvasive and invasive methodologies. For pursuing the origins of language, the significance of imaging and of neurophysiological investigations with intracortical methods in nonhuman animals cannot be easily overestimated. It is indeed likely to facilitate a deeper understanding of the human imaging data and fathom the neural processes and cellular mechanisms that underlie our linguistic capacities. Behaviorally, we know that nonhuman animals have simpler verbal communication systems, so nobody would expect to find the exact components of human speech and language processing in nonhuman animals. Instead, however, as we have illustrated here, the comparative method is most useful when it reveals both the cross-species similarities and the differences in brain structure and function. Many outstanding issues remain (for some, see Box 1), but we can reasonably expect that the comparative approach will reveal further similarities to clarify what cognitive abilities we can model in nonhuman animals for the study of normal and impaired processes at the neuronal level. At the same time, the differences in structure and function will reveal the evolutionary specialization of brain function that has occurred.

Note

1. Like the processing of facial information, voice processing need not contain linguistic information. Yet, it is possible that human vocal and verbal processing are intimately related, just as the verbal source (voice) and its content (speech) often are. As examples, some aspects of speech content are linked to voice qualities, such as sentence prosody, which is largely mediated through changes in voice pitch. Also, voice quality needs to be "normalized" for us to understand spoken language regardless of who is speaking it (e.g., Smith and others 2005), and the human voice regions are strongly activated by speech sounds or neighbor speech sensitive regions of the brain (see Fig. 2*a*; and for a discussion of such specificity issues, Price and others 2005).

References

Alain C, Arnott SR, Hevenor S, Graham S, Grady CL. 2001. "What" and "where" in the human auditory system. Proc Natl Acad Sci U S A 98:12301–6.

Amunts K, Schleicher A, Burgel U, Mohlberg H, Uylings HB, Zilles K. 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J Comp Neurol 412:319–41.

Bates E, Wilson SM, Saygin AP, Dick F, Sereno MI, Knight RT, and others. 2003. Voxel-based lesion-symptom mapping. Nat Neurosci 6:448–50.

Belin P, Zatorre RJ. 2003. Adaptation to speaker's voice in right anterior temporal lobe. Neuroreport 14:2105–9.

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. 2000. Voice-selective areas in human auditory cortex. Nature 403:309–12.

Broca P. 1861. Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). Bulletin de la Société d'Anatomie 6:330–57.

Damasio AR, Geschwind N. 1984. The neural basis of language. Annu Rev Neurosci 7:127–47.

Dehaene-Lambertz G, Pallier C, Serniclaes W, Sprenger-Charolles L, Jobert A, Dehaene S. 2005. Neural correlates of switching from auditory to speech perception. Neuroimage 24:21–33.

Dronkers NF. 1996. A new brain region for coordinating speech articulation. Nature 384:159–61.

Dronkers NF, Plaisant O, Iba-Zizen MT, Cabanis EA. 2007. Paul Broca's historic cases: high resolution MR imaging of the brains of Leborgne and Lelong. Brain 130(Pt 5):1432–41.

Eggert GH, editor. 1977. Wernicke's works on aphasia: a sourcebook and review. The Hague: Mouton, 1977.

Fitch WT, Hauser MD. 2004. Computational constraints on syntactic processing in a nonhuman primate. Science 303:377–80.

Friederici AD, Bahlmann J, Heim S, Schubotz RI, Anwander A. 2006. The brain differentiates human and non-human grammars: functional localization and structural connectivity. Proc Natl Acad Sci U S A 103:2458–63.

Friederici AD, Kotz SA. 2003. The brain basis of syntactic processes: functional imaging and lesion studies. Neuroimage 20(Suppl 1):S8–17.

Fullerton BC, Pandya DN. 2007. Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. J Comp Neurol 504:470–98.

Funkenstein HH, Winter P. 1973. Responses to acoustic stimuli of units in the auditory cortex of awake squirrel monkeys. Exp Brain Res 18:464–88.

Ghazanfar AA, Hauser MD. 2001. The auditory behaviour of primates: a neuroethological perspective. Curr Opin Neurobiol 11:712–20.

Ghazanfar AA, Schroeder CE. 2006. Is neocortex essentially multisensory? Trends Cogn Sci 10:278–85.

Gifford GW 3rd, MacLean KA, Hauser MD, Cohen YE. 2005. The neurophysiology of functionally meaningful categories: macaque ventrolateral prefrontal cortex plays a critical role in spontaneous categorization of species-specific vocalizations. J Cogn Neurosci 17:1471–82.

Gil-da-Costa R, Braun A, Lopes M, Hauser MD, Carson RE, Herscovitch P, and others. 2004. Toward an evolutionary perspective on conceptual representation: species-specific calls activate visual and affective processing systems in the macaque. Proc Natl Acad Sci U S A 101:17516–21.

Gil-da-Costa R, Martin A, Lopes MA, Munoz M, Fritz JB, Braun AR. 2006. Species-specific calls activate homologs of Broca's and Wernicke's areas in the macaque. Nat Neurosci 9:1064–70.

Griffiths TD, Warren JD, Scott SK, Nelken I, King AJ. 2004. Cortical processing of complex sound: a way forward? Trends Neurosci 27:181–5.

Hauser MD, Chomsky N, Fitch WT. 2002. The faculty of language: what is it, who has it, and how did it evolve? Science 298:1569–79.

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

Jarvis ED. 2004. Learned birdsong and the neurobiology of human language. Ann NY Acad Sci 1016:749–77.

Jarvis ED, Gunturkun O, Bruce L, Csillag A, Karten H, Kuenzel W, and others. 2005. Avian brains and a new understanding of vertebrate brain evolution. Nat Rev Neurosci 6:151–9.

Keliris GA, Shmuel A, Ku SP, Pfeuffer J, Oeltermann A, Steudel T, and others. 2007. Robust controlled functional MRI in alert monkeys at high magnetic field: effects of jaw and body movements. Neuroimage 36:550–70.

Liebenthal E, Binder JR, Spitzer SM, Possing ET, Medler DA. 2005. Neural substrates of phonemic perception. Cereb Cortex 15:1621–31.

Lomber SG, Malhotra S. 2008. Double dissociation of 'what' and 'where' processing in auditory cortex. Nat Neurosci 11:609–16.

Marslen-Wilson WD, Tyler LK. 2007. Morphology, language and the brain: the decompositional substrate for language comprehension. Philos Trans R Soc Lond B Biol Sci 362:823–36.

Naeser MA, Alexander MP, Helm-Estabrooks N, Levine HL, Laughlin SA, Geschwind N. 1982. Aphasia with predominantly subcortical lesion sites: description of three capsular/putaminal aphasia syndromes. Arch Neurol 39:2–14.

Obleser J, Boecker H, Drzezga A, Haslinger B, Hennenlotter A, Roettinger M, and others. 2006. Vowel sound extraction in anterior superior temporal cortex. Hum Brain Mapp 27:562–71.

Obleser J, Zimmermann J, Van Meter J, Rauschecker JP. 2007. Multiple stages of auditory speech perception reflected in event-related FMRI. Cereb Cortex 17: 2251–7.

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. 1988. Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature 331:585–9.

Petkov CI, Kayser C, Augath M, Logothetis NK. 2006. Functional imaging reveals numerous fields in the monkey auditory cortex. PLoS Biol 4:e215.

Petkov CI, Kayser C, Steudel T, Whittingstall K, Augath M, Logothetis NK. 2008. A voice region in the monkey brain. Nat Neurosci 11:367–74.

Petrides M, Pandya DN. 2002. Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. Eur J Neurosci 16:291–310.

Poremba A, Malloy M, Saunders RC, Carson RE, Herscovitch P, Mishkin M. 2004. Species-specific calls evoke asymmetric activity in the monkey's temporal poles. Nature 427:448–51.

Poremba A, Saunders RC, Crane AM, Cook M, Sokoloff L, Mishkin M. 2003. Functional mapping of the primate auditory system. Science 299:568–72.

Price C, Thierry G, Griffiths T. 2005. Speech-specific auditory processing: where is it? Trends Cogn Sci 9:271–6.

Rauschecker JP. 1998. Parallel processing in the auditory cortex of primates. Audiol Neurootol 3:86–103.

Rauschecker JP, Tian B, Hauser M. 1995. Processing of complex sounds in the macaque nonprimary auditory cortex. Science 268:111–4.

Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, and others. 2008. The evolution of the arcuate fasciculus revealed with comparative DTI. Nat Neurosci 11:426–8.

Rimol LM, Specht K, Weis S, Savoy R, Hugdahl K. 2005. Processing of sub-syllabic speech units in the posterior temporal lobe: an fMRI study. Neuroimage 26:1059–67.

Rizzolatti G, Arbib MA. 1998. Language within our grasp. Trends Neurosci 21:188–94.

Romanski LM, Averbeck BB, Diltz M. 2005. Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. J Neurophysiol 93:734–47.

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. 1999. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 2:1131–6.

Russ BE, Ackelson AL, Baker AE, Cohen YE. 2008. Coding of auditory-stimulus identity in the auditory non-spatial processing stream. J Neurophysiol 99:87–95.

Schmahmann JD, Pandya DN, Wang R, Dai G, D'Arceuil HE, de Crespigny AJ, and others. 2007. Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography. Brain 130(Pt3): 630–53.

Scott SK. 2005. Auditory processing—speech, space and auditory objects. Curr Opin Neurobiol 15:197–201.

Scott SK, Blank CC, Rosen S, Wise RJ. 2000. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123(Pt 12):2400–6.

Scott SK, Johnsrude IS. 2003. The neuroanatomical and functional organization of speech perception. Trends Neurosci 26:100–7.

Smith DR, Patterson RD, Turner R, Kawahara H, Irino T. 2005. The processing and perception of size information in speech sounds. J Acoust Soc Am 117:305–18.

Suga N. 1992. Philosophy and stimulus design for neuroethology of complex-sound processing. Philos Trans R Soc Lond B Biol Sci 336:423–8.

Suga N, O'Neill WE, Manabe T. 1979. Harmonic-sensitive neurons in the auditory cortex of the mustache bat. Science 203:270–4.

Thompson-Schill SL, D'Esposito M, Aguirre GK, Farah MJ. 1997. Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. Proc Natl Acad Sci U S A 94:14792–7.

Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. 2001. Functional specialization in rhesus monkey auditory cortex. Science 292:290–3.

Ungerleider LG, Mishkin M. 1982. Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJW, editors. Analysis of visual behavior. Cambridge, MA: MIT Press. p 549–80.

von Kriegstein K, Eger E, Kleinschmidt A, Giraud AL. 2003. Modulation of neural responses to speech by directing attention to voices or verbal content. Brain Res Cogn Brain Res 17:48–55.

Wang X, Kadia SC. 2001. Differential representation of species-specific primate vocalizations in the auditory cortices of marmoset and cat. J Neurophysiol 86:2616–20.

Warren JE, Wise RJ, Warren JD. 2005. Sounds do-able: auditory-motor transformations and the posterior temporal plane. Trends Neurosci 28:636–43.

Wernicke C. 1874. Der aphasische Symptomenkomplex. In: Eggert GH, editor. Wernicke's works on aphasia: a sourcebook and review. The Hague: Mouton, 1977. p 91–145.

Wilson SM, Saygin AP, Sereno MI, Iacoboni M. 2004. Listening to speech activates motor areas involved in speech production. Nat Neurosci 7:701–2.

Wise RJ, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA. 2001. Separate neural subsystems within 'Wernicke's area.' Brain 124(Pt 1):83–95.
