Published in final edited form as:
Hear Res. 2011 January ; 271(1-2): 16–25. doi:10.1016/j.heares.2010.09.001.

An Expanded Role for the Dorsal Auditory Pathway in Sensorimotor Control and Integration

Josef P. Rauschecker
Department of Physiology and Biophysics, Georgetown University School of Medicine, Washington, DC 20057-1460, U.S.A

Abstract

The dual-pathway model of auditory cortical processing assumes that two largely segregated processing streams originating in the lateral belt subserve the two main functions of hearing: identification of auditory "objects", including speech; and localization of sounds in space (Rauschecker and Tian, 2000). Evidence has accumulated, chiefly from work in humans and nonhuman primates, that an antero-ventral pathway supports the former function, whereas a postero-dorsal stream supports the latter, i.e. processing of space and motion-in-space. In addition, the postero-dorsal stream has also been postulated to subserve some functions of speech and language in humans. A recent review (Rauschecker and Scott, 2009) has proposed the possibility that both functions of the postero-dorsal pathway can be subsumed under the same structural forward model: an efference copy sent from prefrontal and premotor cortex provides the basis for "optimal state estimation" in the inferior parietal lobe and in sensory areas of the posterior auditory cortex. The current article corroborates this model by adding and discussing recent evidence.

Keywords

Auditory cortex; dual-pathway model; posterior-dorsal stream; space; motion; speech; Wernicke; phonological-articulatory loop; efference copy; internal models; forward models; optimal state estimation

Introduction: Dual Processing Streams in Vision and Audition

Throughout evolution, all sensory systems - and vision and audition in particular - have served dual functions for perception and behavior: identification of sensory stimuli or events and localization of these stimuli in space. In the cerebral cortex of mammals, largely segregated anatomical pathways can be discerned for both vision and audition. These pathways originate in the respective primary cortical areas and display a division of labor with regard to identification and localization (Rauschecker and Tian, 2000; Ungerleider and Mishkin, 1982). In the auditory cortical system, areas anterior and lateral to auditory core form a hierarchical processing stream which ultimately leads to the storage and recognition

---

Address for correspondence: J.P. Rauschecker, PhD, Laboratory of Integrative Neuroscience and Cognition, Georgetown University Medical Center, New Research Building, Room WP19, Washington, DC 20057-1460, USA, Tel.: ++1-202-687-8842, Fax: ++1-202-687-0617, rauschej@georgetown.edu.

of specific feature combinations. Some authors have referred to these entities as "objects" (Griffiths and Warren, 2004; Zatorre et al., 2004), a term that has more intuitive meaning in vision but may be explained as short hand for: "specific feature combinations used for the identification of a stimulus". The role of the antero-ventral auditory processing stream in auditory object identification, or as a "what"-pathway, has become largely undisputed (Leaver and Rauschecker, 2010; Rauschecker and Scott, 2009). Amongst this pathway's functions is the decoding of species-specific communication sounds in animals or speech in humans.

A second pathway originating from auditory cortical core areas projects posteriorly and dorsally. The present article will focus on this second pathway, which has traditionally been defined as a spatial or "where"-pathway, equivalent to the dorsal pathway in vision. Much of the evidence reviewed here, from animals as well as humans, supports such a role of the postero-dorsal auditory pathway for spatial hearing. However, results from human studies, both classical patient work (Galaburda, 1993; Geschwind, 1965) and modern neuroimaging data (Hickok and Poeppel, 2000; Hickok and Poeppel, 2007), have also pointed to a role of the dorsal auditory pathway in speech and language, especially in terms of a phonological or articulatory loop (Baddeley et al., 1984). The seeming incompatibility of these two distinct bodies of data (space versus speech) has led to calls for giving up the concept of dual processing streams altogether. Others have called for an abandonment of the human-monkey comparison due to supposedly fundamental species differences: whereas monkeys use their dorsal stream for space processing, humans use it for speech (Hickok and Poeppel, 2007). As we have argued recently (Rauschecker and Scott, 2009) and will reiterate here, the neural functions related to space and speech, in a computational sense, may not be as incompatible as they seem. Rather, both share a common set of properties that actually require a neural system like the dorsal stream, which creates an interface between sensory and motor networks and performs a matching operation between predicted outcomes and actual events. While the actual computational algorithms in the brain are far from clear, they must resemble the internal "forward models" that have revolutionized thinking in motor control and robotics (Kawato, 1999; Wolpert et al., 1995).

This expanded concept of the dorsal stream not only unifies sensorimotor aspects of space and speech within the auditory domain; it also generalizes dorsal-stream function between vision and audition. In doing so, the revised concept turns some of the conventional wisdom about the dorsal stream on its head: it transforms it from a purely sensory or afferent pathway into an equally efferent pathway, in which predictive motor signals modify activity in sensory structures. As such, the present theory obviates the postulate for a third pathway (for "how" or "when") (Battelli et al., 2008; Belin and Zatorre, 2000; Schubotz et al., 2003; Scott, 2005; Spierer et al., 2009a), as aspects of that are incorporated in the current dual-pathway concept.

Auditory Space Processing in the Dorsal Pathway of Cats and Monkeys

Spatial tuning of cortical neurons

Although it is common knowledge that brainstem mechanisms play an important role in the processing of spatial attributes of sounds (Irvine, 1992; King and Nelken, 2009; Knudsen and Konishi, 1978), early studies have also suggested a role for auditory cortex in sound localization (Diamond et al., 1956; Heffner and Masterton, 1975; Ravizza and Masterton, 1972).

In rhesus monkeys, core areas A1 and R (the "primary" and "rostral" fields, respectively), are surrounded by secondary belt areas (Kaas and Hackett, 2000). Both lateral and medial belt (LB and MB) neurons respond better to band-passed noise bursts than to pure tones

(Kusmierek and Rauschecker, 2009; Rauschecker et al., 1995). Comparing core and belt, spatially tuned neurons are present in A1 but are found at a much higher density in the caudo-medial belt field (CM) (Rauschecker et al., 1997; Recanzone, 2000). When monkeys are trained in an auditory localization task, the firing rate of neurons in CM correlates more tightly with behavioral performance than that of neurons in A1, which is a strong indication that CM plays an important role in sound localization (Recanzone et al., 2000). Such localization is most likely accomplished on the basis of a population code (Miller and Recanzone, 2009).

To directly compare spatial selectivity of neurons in the rostral and caudal LB in the same animals, broad-band species-specific communication calls were presented from different locations (Tian et al., 2001). The highest spatial selectivity was found in the caudolateral (CL) and the lowest in the anterolateral area (AL). Together with the connectivity studies described below, this has led to the hypothesis that the caudal belt forms the beginning of a cortical processing stream for auditory space, whereas AL forms the beginning of a non-spatial “what”-stream (Rauschecker and Tian, 2000; Tian et al., 2001). The middle lateral area (ML) does not seem specialized for either of these functions, which fits with the fact that it appears to be at a lower hierarchical level anatomically compared to both AL and CL. The latter is suggested by anatomical tracer studies that show input from the main relay nucleus of the auditory thalamus (MGv) to area ML (in addition to core areas), but not to AL and CL (Rauschecker, 1998; Zhenochin et al., 1998) [c.f. (Hackett et al., 1998b) and (Morel et al., 1993), who point out that A1 and “cortex immediately lateral to A1” have tonotopically organized connections with MGv]. Furthermore, ML stains more darkly for parvalbumin, myelin, and acetylcholinesterase than, for instance, AL (Hackett et al., 1998a), which is also an indication that it is more core-like.

Auditory projections to prefrontal and parietal cortex

Connectivity studies in rhesus monkeys (Rauschecker et al., 1997) have shown that at least one of the caudal belt regions receives its subcortical input via a separate pathway than core areas A1 and R. While the latter receive projections from the principal relay nucleus of the auditory thalamus, the ventral nucleus of the medial geniculate (MGv), area CM receives projections from its dorsal and medial subnuclei (MGd and MGm). This parallel input pathway to areas of the supratemporal plane may start even earlier: Single-unit studies indicate that the dorsal cochlear nucleus (DCN) has response properties compatible with a function in auditory space processing (May, 2000). Thus, area CM could receive at least some of its input from the DCN via the external nuclei of the inferior colliculus and the MGd (Rauschecker, 1997), although interaural timing cues are also relayed via the ventral cochlear nucleus (VCN).

Anatomical tracer studies have demonstrated the existence of largely segregated pathways that originate in the LB and project to different target regions in the prefrontal cortex (Romanski et al., 1999). Injections into CL specifically led to labeling of dorsolateral prefrontal cortex (DLPFC; areas 8a, 46), which is known for its involvement in spatial working memory (Goldman-Rakic, 1996). Conversely, injections into AL led to labeling of ventrolateral prefrontal cortex (VLPFC). As one might expect, neurons in VLPFC were reliably modulated during a non-spatial auditory task but were not modulated during a spatial auditory task (Cohen et al., 2009). Although recent neuroanatomical studies have added much detail to this model of connectivity in the auditory cortex, the overall thrust of the earlier work has held up: starting out in core areas, two main directions of anatomical projections can be discerned: an anterior and a posterior processing stream (Hackett, 2010; Romanski and Averbeck, 2009).

A projection from posterior STG (caudal belt) to posterior parietal (PP) cortex in monkeys was found independently by Lewis and Van Essen (2000). Specifically, the ventral intraparietal area (VIP) was identified as the primary recipient of auditory input to PP. Activation of inferior parietal lobule (IPL) by sound localization has also been demonstrated in human imaging studies (Bushara et al., 1999; Griffiths et al., 1998; Maeder et al., 2001; Weeks et al., 1999), as discussed in greater detail in the following section. A role of the IPL in the processing of auditory space is also evident from human clinical studies (Clarke et al., 2000; Griffiths et al., 1996; Griffiths et al., 1997). Some findings are reminiscent of the phenomenon of spatial neglect first described in the visual modality after lesions of right parietal cortex. This suggests that the various sensory modalities are eventually combined into one unitary spatial representation (Spierer et al., 2009b).

Auditory Space Processing in the Dorsal Pathway of Humans

Core, belt and parabelt areas

Using the same types of stimuli as in the preceding monkey studies, human neuroimaging work also suggests an organization of auditory cortex into core, belt and parabelt areas. Two core areas robustly activated by pure-tone stimuli and with mirror-symmetric tonotopic organization were found along Heschl’s gyri (Formisano et al., 2003; Wessinger et al., 2001). A third such area was sometimes seen more laterally. While the first two areas obviously correspond to core areas A1 and R, the third may be homologous to area RT or to ML, which is more primary-like on some accounts than other belt areas (see above). As observed in monkeys, the pure-tone (PT) responsive areas were surrounded by belt regions both medially and laterally, which were activated only by BPN bursts (Wessinger et al., 2001). Although the study of tonotopic organization in human auditory cortex has remained a vexing problem (Humphries et al., 2010), recent data from our lab corroborate a subdivision into core, belt and parabelt areas in human auditory cortex based on responses to PT, BPN and vowel sounds (Chevillet et al., 2010).

Role of dorsal auditory stream in spatial processing

Antero-lateral areas of the superior temporal cortex are activated by complex natural sounds of a non-spatial nature as well as intelligible speech or speech-like sounds (Alain et al., 2001; Binder et al., 2004; Binder et al., 2000; Leaver and Rauschecker, 2010; Maeder et al., 2001; Obleser et al., 2006; Scott et al., 2000). Thus it appears likely that behaviorally relevant auditory objects, including speech sounds, are identified within an anterior-lateral auditory “what”-stream.

Auditory areas located in the “planum temporale” (PT) posterior to Heschl’s gyrus are less selective for auditory object categories and seem to be involved in a variety of auditory functions (Smith et al., 2009), including the processing of music (Hyde et al., 2008; Zatorre et al., 2002a). A wider role of PT and posterior STG for processing spectro-temporally complex sounds has therefore been postulated (Belin et al., 2000; Nourski et al., 2009; Obleser et al., 2007). In its most general form it has led to the suggestion of PT as a “computational hub” (Griffiths and Warren, 2002).

Further posterior in the STG and STS are regions of the caudal belt and parabelt (projecting up dorsally into the inferior posterior parietal cortex) that are activated during spatial tasks, such as auditory spatial discrimination or tasks involving auditory motion in space (Arnott et al., 2004; Brunetti et al., 2005; Degerman et al., 2006; Jääskeläinen et al., 2004; Krumbholz et al., 2005a; Krumbholz et al., 2005b; Maeder et al., 2001; Tata and Ward, 2005a; Tata and Ward, 2005b; Warren et al., 2002; Zatorre et al., 2002b; Zimmer and Macaluso, 2005).

In a PET study by Zatorre et al. (2002b), the posterior auditory cortex responded to sounds varying in spatial distribution, but only when multiple complex stimuli were presented simultaneously. These authors also found the right inferior parietal cortex to be recruited specifically during localization tasks, which is consistent with other studies, e.g. (Griffiths et al., 1998). An fMRI study by Krumbholz and co-workers (2005b) found that interaural time differences were represented along a posterior pathway comprising the planum temporale (PT) and IPL of the respective contralateral hemisphere. In contrast to Zatorre et al. (2002b), this study found that compared to a centrally presented sound, stationary lateralized sounds did produce a significant activation increase in the PT of the respective contralateral hemisphere.

In the study of Krumbholz et al. (2005b) the response was stronger (and extended further into adjacent regions of the IPL) when the sound was moving than when it was stationary, which is a finding that conforms to earlier results by Warren et al. (2002). Sounds moving in space are vastly more complex in computational terms than visual stimuli moving in space: not only must the spatial positions of these sounds be computed via binaural and/or monaural cues, but also these representations must be updated on a moment-by-moment basis in order to extract movement information. It is conceivable that auditory spatial representations that can be combined with spatial representations in other modalities do exist.

Timing differences between the two ears can be used to localize sounds in space only when the inputs to the two ears have similar spectro-temporal profiles (high binaural coherence). Zimmer and Macaluso (2005) found that activity in Heschl’s gyrus increased with increasing coherence, irrespective of localization being task-relevant. Posterior auditory regions also showed increased activity for high coherence, but only when sound localization was required and subjects successfully localized sounds. The authors concluded that binaural coherence cues are processed throughout the auditory cortex but that these cues are specifically used by posterior regions of the STG for successful auditory localization (Zimmer and Macaluso, 2005). In another series of fMRI experiments, Deouell et al. (2007) showed that a region in the human medial PT is sensitive to auditory spatial changes, even when subjects are not engaged in a sound localization task, i.e, when the spatial changes are occurring in the background. Thus, acoustic space is firmly represented in the human PT even when sound processing is not required by the ongoing task.

Tata and Ward (2005a; 2005b) used auditory evoked potentials to explore the putative auditory “where”-pathway in humans. The mismatch negativity (MMN) elicited by deviations in sound location is comprised of temporally and anatomically distinct phases: an early phase with a generator posterior to primary auditory cortex and contralateral to the deviant stimulus, and a later phase with generators that are more frontal and bilaterally symmetric. The posterior location of the early-phase generator suggests the engagement of neurons within a posterior “where”-pathway for processing spatial auditory information (Tata and Ward, 2005a).

In a study combining fMRI and magnetoencephalography (MEG), Brunetti and co-workers found that the processing of sound coming from different locations activates a neural circuit similar to the auditory “where” pathway described in monkeys (Brunetti et al., 2005). This system included Heschl’s gyrus, the posterior STG, and the inferior parietal lobule. MEG analysis enabled the timing of this circuit to be assessed: activation of Heschl’s gyrus was observed 139 ms after the auditory stimulus, the peak latency of the source located in the posterior STG occurred at 156 ms, and the inferior parietal lobule and the supramarginal gyrus peaked at 162 ms. Both hemispheres were involved in the processing of sounds originating from different locations, but a stronger activation was observed in the right hemisphere (Brunetti et al., 2005).

A similar study combining fMRI and MEG was conducted by Ahveninen et al. (2006). They found a double dissociation in response adaptation to sound pairs with phonetic vs. spatial sound changes, demonstrating that the human nonprimary auditory cortex indeed processes speech-sound identity and location in parallel anterior “what” (in anterolateral Heschl’s gyrus, anterior superior temporal gyrus, and posterior planum polare) and posterior “where” (in PT and posterior STG) pathways as early as approximately 70–150 ms after stimulus onset. These data further showed that the “where” pathway is activated approximately 30 ms earlier than the “what” pathway.

Even before some of the latest and most conclusive studies were published, Arnott et al. (2004), in a meta-analysis, reviewed evidence from auditory functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) studies to determine the reliability of the auditory dual-pathway model in humans. Activation coordinates from 11 “spatial” studies (i.e., listeners made localization judgments on sounds that could occur at two or more perceptually different positions) and 27 “nonspatial” studies (i.e., listeners completed nonspatial tasks involving sounds presented from the same location) were entered into the analysis. Temporal lobe activity during spatial tasks was confined almost exclusively to posterior areas. In addition, all but one of the spatial studies reported activation within the IPL as opposed to only 41% of the nonspatial studies. Inferior frontal activity (Brodmann’s areas 45 and 47) was reported in only 9% of the spatial dual studies, but in 56% of the nonspatial studies.

These results support a model in which nonspatial sound information (e.g., sound identity) is processed primarily along an antero-ventral stream whereas sound location and motion in space are processed exclusively along a postero-dorsal stream, i.e. within auditory areas posterior to the primary auditory cortex and in the parietal cortex, projecting to DLPFC. Furthermore, it appears that, as in the visual system, studies of nonhuman primates can serve as excellent models for human studies and a major species difference need not be claimed on these grounds. Conversely, human imaging studies can provide useful guidance for microelectrode studies in nonhuman primates, which permit analyses at higher spatial and temporal resolution.

Role of the Human Dorsal Auditory Pathway in Speech and Language

Is the dorsal pathway really involved in speech perception?

In the preceding section I have summarized the evidence for a role of the posterior ST (pST) region (and the IPL regions connected with it) in processing auditory space and motion-in-space. This function is undeniably present in both monkeys and humans (as well as non-primate animals). However, another view about the function of pST in humans has classically been even more widespread: the view that pST is involved in speech perception or comprehension (Damasio and Damasio, 1980; Geschwind, 1965). Many textbooks refer to pST as “Wernicke’s area”, so it seems as if this view dates back to Carl Wernicke (1874), who described patients with lesions of the ST region having difficulties with various aspects of speech. Closer examination of Wernicke’s case studies reveals, however, that the pertinent lesions were not necessarily found in pST alone. A figure in one of his own textbooks (Wernicke, 1881) explicitly marked the whole ST region as speech-related, including its anterior aspects. To reserve the term “Wernicke’s area” for the posterior one-third of ST is, therefore, misleading.

Wernicke did, however, make the insightful claim that auditory ST regions subserving the deciphering of speech sounds must be connected somehow with the motor speech area in the frontal cortex, which had been discovered by Broca (1861) about a decade earlier. Based on gross anatomical studies of aphasic stroke patients, later researchers assumed that this functional connectivity was provided by a fiber bundle that wound its way from the posterior ST region to Broca's area, the "arcuate fascicle" (Geschwind, 1965). Present-day work is being performed with high-resolution structural imaging techniques {Bernal, 2009 #890; Keller et al., 2009; Rilling et al., 2008). At least one of these studies has revealed that a direct connection from pST to Broca's area, as in the monkey and its homologous areas (Petrides and Pandya, 2009), barely exists (Frey et al., 2008). Instead, most fibers projecting to Broca's area from ST originate in its anterior aspects and follow a whole different pathway via the extreme capsule and/or the uncinate fascicle (Ebeling and von Cramon, 1992; Friederici et al., 2006)). In fact, Wernicke himself suspected that the connection from ST to Broca's area went via the anterior insula, a region that has recently been found to play a role in communication sound processing of monkeys (Remedios et al., 2009). All this adds to the support for an antero-ventral pathway in auditory speech processing and one might be tempted to reject the claim of a specific pST (and dorsal-stream) involvement in speech processing altogether. However, this would be throwing the baby out with the bathwater.

In order to salvage a genuine role for the pST region in speech and language and to reconcile this role with the spatial functions of that region, one merely has to back away from the claim that pST is involved in the "perception" of speech, that is, primarily an acoustic-phonetic decoding of speech sounds. Instead, one needs to analyze the incidents under which pST areas and parietal cortex in the IPL are activated by sounds or tasks with other than spatial connotations.

Representation of action sounds in the dorsal stream

Various studies have demonstrated activation of left parietal cortical regions while subjects were listening to sounds generated by actions, such as tool sounds (Engel et al., 2009; Lewis et al., 2005; Pizzamiglio et al., 2005). These activations often include posterior STS and STG regions, especially when contrasted with unrecognizable control sounds. One possibility is that these regions contain representations of "doable" sounds (Rauschecker and Scott, 2009). In particular, it has been suggested that the medial PT region (Warren et al., 2005) contains templates of "doable" articulations (not limited to speech sounds) against which incoming sounds are matched. Studies of silent articulation (Wise et al., 2001) and covert rehearsal of speech (Hickok et al., 2009; Hickok et al., 2000) have also identified activation in the posterior medial PT region within the posterior-dorsal stream.

Such findings resonate with the "affordance" model of Gibson (1977), where objects and events are described in terms of action possibilities. Gibson's views undoubtedly had an influence on the mirror-neuron theory of Rizzolatti and colleagues (2006). More specifically with regard to speech, the above findings are reminiscent of the "motor theory of speech perception" (Liberman and Mattingly, 1985). Interestingly, several regions of the posterior-lateral-temporal cortex (PLTC) are also reliably recruited when participants read or listen to action verbs (Bedny et al., 2008; Reale et al., 2007), thus reaching into the realm of abstract concepts.

A multisensory reference frame

The postero-medial region of the PT has been identified as a possible key node for the feedback control of speech production (Dhanjal et al., 2008) since it shows a response to somatosensory input from articulators as well as to auditory speech input. Adjacent to pST, the temporo-parietal junction (TPJ) has been discussed independently in both auditory and visual contexts, but probably constitutes a multisensory region having to do with temporal order judgment of spatially separate events (Davis et al., 2009).

In relation to these studies, it is fitting that neurophysiological evidence from nonhuman primates shows that auditory caudal belt areas are not only responsive to auditory input but reveal multisensory responses (Brosch et al., 2005; Fu et al., 2003; Ghazanfar et al., 2005; Kayser et al., 2007; Lakatos et al., 2007). Neuroanatomical studies demonstrate that both caudal medial and lateral belt fields receive input from somatosensory and multisensory cortex as well as thalamic nuclei (Smiley et al., 2007). In contrast, core and anterior areas show only sparse multisensory connections. Thus, the posterior-dorsal stream, by bringing together input from different sensory modalities, may create a supramodal reference frame in which any transformations, whether spatial or otherwise, can be conducted.

Encoding and retrieval of sound sequences

One of the unsolved puzzles in auditory neuroscience is how the brain encodes and stores sequences of sound (Rauschecker, 2005; Schubotz et al., 2000). Unlike tape recorders and compact disk players the brain does not have any moving parts that could translate temporal order of a sound sequence into location on a physical medium for storage and retrieval. Digital music players, on the other hand, use specific file formats to preserve the spectro-temporal integrity of, for instance, a piece of music. If we look for structures in the brain that may be suitable for storage and reproduction of temporal sequences, we are quickly reminded of the fact that motor areas must be able to do just that: a simple motor act or gesture requires the production of sequences of nerve signals sent to specific muscles (or motor neurons) controlling the various limbs involved in that gesture in a particular order. The act of speaking or singing is an example of a motor performance during which a multitude of fine-grained muscles have to be controlled in a highly time-order specific fashion in order to keep both rhythm and pitch exactly right. While the motor cortex provides the origin of axons projecting to the spinal cord for control of muscles, it is commonly assumed that subcortical entities such as the basal ganglia or the cerebellum set up the patterns reflecting temporal sequential structure of motor acts.

Indeed, singing or speaking, like other motor acts, light up cortical motor areas as well as subcortical structures (Perry et al., 1999). Singing also activates auditory areas, which would not be surprising (because the subjects hear their own voice) if the activation didn’t persist even after subtracting out auditory perceptual activation. Interestingly, the remaining auditory activation appears in pST. Even more interestingly, listening to music also activates motor areas (Chen et al., 2008; Wilson et al., 2004; Zatorre et al., 2007). It thus appears as if we are looking at a sensorimotor loop, wherein both afferent and efferent branches are active in either situation.

Finally, even imagery of music (Halpern and Zatorre, 1999) and anticipation of familiar melodies after playing the preceding melody (Leaver et al., 2009) leads to activation of both auditory and motor structures (Fig. 1), cortical and subcortical (cerebellum and basal ganglia). The amount of basal ganglia versus frontal cortical activation depends on the state of familiarity of the sequence with basal ganglia more active during the learning period (Leaver et al., 2009).

There is also strong psychophysical evidence suggesting that auditory-motor processing dissociates from auditory-perceptual processing (Rauschecker and Scott, 2009; Repp, 2005): Listeners can accurately tap along to auditory sequences, and their motor responses can track changes in the rates of these sequences. This tracking of sequences could occur in the dorsal stream. Functional imaging evidence does indeed suggest that the intraparietal sulcus plays a role in streaming, sequence detection, and dissociation of figure from ground

(Cusack, 2005). These results from human psychophysical and imaging studies would merit further examination in monkey single-unit studies to get at the exact neurophysiological mechanisms of auditory sequence processing and stream segregation (Micheyl et al., 2005).

Auditory perception/production links in voice and speech

Monkey studies have shown that neurons in auditory cortex are suppressed during vocalization (Eliades and Wang, 2003; Müller-Preuss and Ploog, 1981). This finding is consistent with results from humans, which indicate that superior temporal areas are suppressed during speech production (Curio et al., 2000; Houde et al., 2002; Numminen et al., 1999; Paus et al., 1996). This suppression or attenuation of auditory cortex is found even with covert articulation and lipreading, suggesting the existence of an efference-copy pathway from premotor regions to auditory cortex (Kauramäki et al., 2010) (Fig. 2).

It has been argued that mechanisms of this kind may exist to help distinguish the effects of actions caused by oneself from those caused by the actions of others (Blakemore et al., 1998), specifically differentiating between one’s own voice and the voices of others (Rauschecker and Scott, 2009). However in nonhuman primate studies, auditory neurons that are suppressed during actual vocalizations are often more activated by distorted vocalizations (Eliades and Wang, 2008). This suggests a role for these neurons in the comparison of information from the auditory and motor systems during speech production (Guenther, 2006). Work in humans using distorted feedback of speech production has indeed shown enhanced bilateral activation in pST to distorted feedback, even if it is below the threshold for explicit awareness (Tourville et al., 2008).

There have also been persistent claims for a role of the IPL, i.e. the angular and supramarginal gyri, in phonology (Caplan et al., 1992), particularly an involvement in the “phonological/articulatory loop” (Aboitiz et al., 2006; Baddeley et al., 1984). This has been confirmed in several functional imaging studies, though the precise localization of activity does vary with the type of task used (Buchsbaum and D’Esposito, 2008; Gelfand and Bookheimer, 2003). What seems clear is that the IPL, like pST, is not driven by acoustic-phonetic factors in speech processing but is associated with more domain-general factors (Friederici et al., 2006; Rauschecker and Scott, 2009).

New work using DTI in humans demonstrates that there are direct connections between the pars opercularis of Broca’s area (BA44) and the IPL (Bernal and Ardila, 2009; Frey et al., 2008; Saur et al., 2008), but hardly at all with pST, calling into question the notion of a direct connection between “Broca’s” and “Wernicke’s” area, as postulated in most textbooks. In addition, there is the known projection from ventral premotor (vPM) cortex to the IPL (Petrides and Pandya, 1984; Petrides and Pandya, 2009), and connections between parietal cortex and pST are also well known (Seltzer and Pandya, 1994); together, this could form the basis for a feed-forward network between speech production areas and posterior temporal auditory areas (Fig. 3).

Unified Function of the Dorsal Stream: Anticipatory Control of Sensorimotor Events

As this review has documented, posterior ST regions and the IPL participate in the processing of auditory space and motion, and integrate input from several modalities. At the same time, pST and IPL in humans are also involved in the processing and imagery of auditory sequences, including speech and music. Both regions receive input from premotor areas in the dorsal and ventral premotor cortex (PMC). PMC also gets activated during listening to music (Chen et al., 2008; Lahav et al., 2007) and even during musical imagery and anticipation (Leaver et al., 2009). One conclusion is that premotor areas are responsible for assembling the motor patterns for the production of musical sequences (by singing or playing a musical instrument). The sounds being produced activate neuronal assemblies in the auditory cortex, which in turn get matched with the corresponding premotor neurons that helped produce the sounds. Thus, specific sensorimotor networks are established which, together, represent the musical melodies in a quasi-motor code. During learning of musical melodies, which occurs in the same way as learning of motor sequences (Hikosaka et al., 1999), subcortical structures like the basal ganglia and the cerebellum are also active in binding the correct sets of sensory and motor neurons together (Leaver et al., 2009). One prediction would be, therefore, that learning to play a new piece on a musical instrument or, for that matter, learning to play a familiar piece on a new instrument, should result in characteristic changes in premotor representations. The same would be expected when passive listening to complex sounds gets replaced by producing these sounds (“action sounds”).

An analogous process can be assumed to be at work during learning of speech and speech production. Once learned, listening to speech activates the same circuits as during speech production. While it may not strictly be accurate to talk about a “motor code” for speech perception (Liberman et al., 1967), correct speech does require a closing of the loop between perception and production and will lead to coactivation of both networks. The connection between auditory areas in the ST and speech planning areas in the frontal cortex around “Broca’s region”, as postulated by Wernicke, runs through aST and inferior frontal cortex; the loop is closed through PMC via IPL and back to auditory cortex (Fig. 3). Learning to produce new sounds in a foreign language should, therefore, lead to changes in both sensory and motor representations of the corresponding sounds.

Visuomotor sequences are planned and executed in a similar fashion. In most of these cases, spatial information in conjunction with motor signals becomes critical, and this is what parietal cortex is commonly known for (Andersen and Cui, 2009; Colby and Goldberg, 1999). While spatial position is an important variable in both visuo-motor and audio-motor behavior, however, the layout of the fronto-parietal-sensory loop is a more general one having to do with sensorimotor planning and control (Mulliken et al., 2008).

This basic structure is best described by “internal models” or “emulators”, as they are known in motor control theory and robotics (Rauschecker and Scott, 2009). Such models have been used to describe reaching movements or planning of movement trajectories using Kalman filters and Bayesian statistics for optimal state estimation (Desmurget and Grafton, 2000; Kawato, 1999; Sabes, 2000; Simon, 2006). More recently, these models have been used to model perception and imagery as well (Grush, 2004; Wolpert et al., 2003). The inferior parietal cortex appears to provide an ideal interface for feed-forward information from motor preparatory networks in the PFC and PMC to be matched with feedback signals from sensory areas. The goal of the internal model is to minimize the resulting error signal in this process. In some instances, the cerebellum and basal ganglia have also been incorporated into these models (Blakemore et al., 1998).

The feed-forward projection from BA 44 and vPM can be considered the pathway carrying an “efference copy” or “corollary discharge” in the classical sense (Sperry, 1950; Von Holst and Mittelstaedt, 1950), informing the sensory system of planned motor articulations that are about to happen. This signal provides a predictive quality to activity running from frontal areas to the IPL, which therefore anticipates the sensory consequences of action. The feedback signal coming to the IPL from posterior ST, on the other hand, can be considered an “afference copy” (Hershberger, 1976) or reafference with relatively short latencies and high temporal precision (Jääskeläinen et al., 2004; Kauramäki et al., 2010). It can be thought of as a sparse but fast primal sketch of ongoing sensory events (Bar et al., 2006) that are compared with the predictive motor signal in the IPL in real time at every instance. In that sense, both spatial processing and real-time processing of speech and music make use of the same general internal model structures that enable the instantiation of smooth sequential motor behaviors, including visuo-spatial reaching as well as articulation of speech. At the same time, these sensorimotor loops also support the disambiguation of phonological information. Perception (via the ventral stream) and action (via the dorsal stream) operate as a dual system (Goodale and Milner, 1992). These systems not only alternate, but in many cases partially or wholly operate in concert (Indefrey and Levelt, 2004).

Acknowledgments

The present chapter draws from the following prior publications: Rauschecker, 2007; Rauschecker and Scott, 2009. The author's work was supported by grants from the National Institutes of Health (R01 NS052494), the Cognitive Neuroscience Initiative of the National Science Foundation (BCS-0519127), and the NSF PIRE program (OISE-0730255). I would like to thank Priyanka Chablani for help with editing.

List of abbreviations

A1	primary auditory cortex
AES	anterior ectosylvian sulcus
SC	superior colliculus
PAF	posterior auditory field
R	rostral field
LB and MB	lateral and medial belt
CM	caudo-medial belt field
CL	caudolateral area
AL	anterolateral area
ML	middle lateral area
MGv	ventral nucleus of the medial geniculate
MGd	dorsal nucleus of the medial geniculate
MGm	medial nucleus of the medial geniculate
DCN	dorsal cochlear nucleus
VCN	ventral cochlear nucleus
DLPFC	dorsolateral prefrontal cortex
VLPFC	ventrolateral prefrontal cortex
PP	posterior parietal cortex
STG	superior temporal gyrus
STS	superior temporal sulcus
aST	anterior superior temporal
pST	posterior superior temporal
VIP	ventral intraparietal area
IPL	inferior parietal lobule
PT	planum temporale

PET	positron emission tomography
fMRI	functional magnetic resonance imaging
DTI	diffusion tensor imaging
MMN	mismatch negativity
MEG	magnetoencephalography
PLTC	posterior-lateral-temporal cortex
TPJ	temporo-parietal junction
BA	Brodmann area
dPMC, vPMC	dorsal and ventral premotor cortex
vPM	ventral premotor
PMC	premotor cortex
PFC	prefrontal cortex
NSF	National Science Foundation
IFG, SFG	inferior and superior frontal gyrus
pre-SMA	pre-supplementary motor area
CD	compact disk
AC	auditory cortex
IFC	inferior frontal cortex
CS	central sulcus

References

Aboitiz F, Garcia RR, Bosman C, Brunetti E. Cortical memory mechanisms and language origins. Brain Lang 2006;98:40–56. [PubMed: 16481035]

Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämäläinen M, Levänen S, Lin FH, Sams M, Shinn-Cunningham BG, Witzel T, Belliveau JW. Task-modulated “what” and “where” pathways in human auditory cortex. Proc Natl Acad Sci U S A 2006;103:14608–13. [PubMed: 16983092]

Alain C, Arnott SR, Hevenor S, Graham S, Grady CL. “What” and “where” in the human auditory system. Proc Natl Acad Sci U S A 2001;98:12301–6. [PubMed: 11572938]

Andersen RA, Cui H. Intention, action planning, and decision making in parietal-frontal circuits. Neuron 2009;63:568–83. [PubMed: 19755101]

Arnott SR, Binns MA, Grady CL, Alain C. Assessing the auditory dual-pathway model in humans. Neuroimage 2004;22:401–408. [PubMed: 15110033]

Baddeley A, Lewis V, Vallar G. Exploring the articulatory loop. The Quarterly Journal of Experimental Psychology A 1984:233–252.

Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, Hämäläinen MS, Marinkovic K, Schacter DL, Rosen BR, Halgren E. Top-down facilitation of visual recognition. Proc Natl Acad Sci U S A 2006;103:449–54. [PubMed: 16407167]

Battelli L, Walsh V, Pascual-Leone A, Cavanagh P. The ‘when’ parietal pathway explored by lesion studies. Curr Opin Neurobiol 2008;18:120–6. [PubMed: 18708141]

Bedny M, Caramazza A, Grossman E, Pascual-Leone A, Saxe R. Concepts are more than percepts: the case of action verbs. J Neurosci 2008;28:11347–53. [PubMed: 18971476]

Belin P, Zatorre RJ. ‘What’, ‘where’ and ‘how’ in auditory cortex. Nat Neurosci 2000;3:965–6. [PubMed: 11017161]

Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. Nature 2000;403:309–12. [PubMed: 10659849]

Bernal B, Ardila A. The role of the arcuate fasciculus in conduction aphasia. Brain 2009;132:2309–16. [PubMed: 19690094]

Binder JR, Liebenthal E, Possing ET, Medler DA, Ward BD. Neural correlates of sensory and decision processes in auditory object identification. Nat Neurosci 2004;7:295–301. [PubMed: 14966525]

Binder JR, Frost JA, Hammeke TA, Bellgowan PS, Springer JA, Kaufman JN, Possing ET. Human temporal lobe activation by speech and nonspeech sounds. Cereb Cortex 2000;10:512–28. [PubMed: 10847601]

Blakemore SJ, Wolpert DM, Frith CD. Central cancellation of self-produced tickle sensation. Nat Neurosci 1998;1:635–40. [PubMed: 10196573]

Broca P. Remarques sur le siège de la faculté du language articulé: suivies d’une observation d’aphémie (perte de la parole). Bull Soc Anat Paris 1861;6:330–357.

Brosch M, Selezneva E, Scheich H. Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. J Neurosci 2005;25:6797–806. [PubMed: 16033889]

Brunetti M, Belardinelli P, Caulo M, Del Gratta C, Della Penna S, Ferretti A, Lucci G, Moretti A, Pizzella V, Tartaro A, Torquati K, Olivetti Belardinelli M, Romani GL. Human brain activation during passive listening to sounds from different locations: An fMRI and MEG study. Hum Brain Mapp 2005;26:251–61. [PubMed: 15954141]

Buchsbaum BR, D’Esposito M. The search for the phonological store: from loop to convolution. J Cogn Neurosci 2008;20:762–78. [PubMed: 18201133]

Bushara KO, Weeks RA, Ishii K, Catalan MJ, Tian B, Rauschecker JP, Hallett M. Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. Nat Neurosci 1999;2:759–66. [PubMed: 10412067]

Caplan D, Rochon E, Waters GS. Articulatory and phonological determinants of word length effects in span tasks. The Quarterly Journal of Experimental Psychology 1992;45:177–92. [PubMed: 1410554]

Chen JL, Penhune VB, Zatorre RJ. Listening to musical rhythms recruits motor regions of the brain. Cereb Cortex 2008;18:2844–54. [PubMed: 18388350]

Chevillet M, Riesenhuber M, Rauschecker JP. Functional localization of the ventral auditory “what” stream hierarchy. 2010

Clarke S, Bellmann A, Meuli RA, Assal G, Steck AJ. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. Neuropsychologia 2000;38:797–807. [PubMed: 10689055]

Cohen YE, Russ BE, Davis SJ, Baker AE, Ackelson AL, Nitecki R. A functional role for the ventrolateral prefrontal cortex in non-spatial auditory cognition. Proc Natl Acad Sci U S A 2009;106:20045–50. [PubMed: 19897723]

Colby CL, Goldberg ME. Space and attention in parietal cortex. Annu Rev Neurosci 1999;22:319–49. [PubMed: 10202542]

Curio G, Neuloh G, Numminen J, Jousmaki V, Hari R. Speaking modifies voice-evoked activity in the human auditory cortex. Hum Brain Mapp 2000;9:183–91. [PubMed: 10770228]

Cusack R. The intraparietal sulcus and perceptual organization. J Cogn Neurosci 2005;17:641–51. [PubMed: 15829084]

Damasio H, Damasio AR. The anatomical basis of conduction aphasia. Brain 1980;103:337–50. [PubMed: 7397481]

Davis B, Christie J, Rorden C. Temporal order judgments activate temporal parietal junction. J Neurosci 2009;29:3182–8. [PubMed: 19279255]

Degerman A, Rinne T, Salmi J, Salonen O, Alho K. Selective attention to sound location or pitch studied with fMRI. Brain Res 2006;1077:123–134. [PubMed: 16515772]

Deouell LY, Heller AS, Malach R, D’Esposito M, Knight RT. Cerebral responses to change in spatial location of unattended sounds. Neuron 2007;55:985–96. [PubMed: 17880900]

Desmurget M, Grafton S. Forward modeling allows feedback control for fast reaching movements. Trends in Cognitive Sciences 2000;4:423–431. [PubMed: 11058820]

Dhanjal NS, Handunnetthi L, Patel MC, Wise RJ. Perceptual systems controlling speech production. J Neurosci 2008;28:9969–75. [PubMed: 18829954]

Diamond IT, Fisher JF, Neff WD, Yela M. Role of auditory cortex in discrimination requiring localization of sound in space. J Neurophysiol 1956;19:500–512. [PubMed: 13377200]

Ebeling U, von Cramon D. Topography of the uncinate fascicle and adjacent temporal fiber tracts. Acta Neurochir (Wien) 1992;115:143–8. [PubMed: 1605083]

Eliades SJ, Wang X. Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. J Neurophysiol 2003;89:2194–207. [PubMed: 12612021]

Eliades SJ, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 2008;453:1102–6. [PubMed: 18454135]

Engel LR, Frum C, Puce A, Walker NA, Lewis JW. Different categories of living and non-living sound-sources activate distinct cortical networks. Neuroimage 2009;47:1778–91. [PubMed: 19465134]

Formisano E, Kim DS, Di Salle F, van de Moortele PF, Ugurbil K, Goebel R. Mirror-symmetric tonotopic maps in human primary auditory cortex. Neuron 2003;40:859–69. [PubMed: 14622588]

Frey S, Campbell JS, Pike GB, Petrides M. Dissociating the human language pathways with high angular resolution diffusion fiber tractography. J Neurosci 2008;28:11435–44. [PubMed: 18987180]

Friederici AD, Bahlmann J, Heim S, Schubotz RI, Anwander A. The brain differentiates human and non-human grammars: functional localization and structural connectivity. Proc Natl Acad Sci U S A 2006;103:2458–63. [PubMed: 16461904]

Fu KG, Shah AS, Arnold L, Garraghty PE, Smiley J, Hackett TA, Schroeder CE. Auditory cortical neurons respond to somatosensory stimulation. J Neurosci 2003;23:7510–7515. [PubMed: 12930789]

Galaburda AM. The planum temporale. Arch Neurol 1993;50:457. [PubMed: 8489400]

Gelfand JR, Bookheimer SY. Dissociating neural mechanisms of temporal sequencing and processing phonemes. Neuron 2003;38:831–42. [PubMed: 12797966]

Geschwind N. Disconnexion syndromes in animals and man. Brain 1965;88:237–294. 585–644. [PubMed: 5318481]

Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. J Neurosci 2005;25:5004–12. [PubMed: 15901781]

Gibson, JJ. The theory of affordances. In: Shaw, R.; Bransford, J., editors. Perceiving, Acting, and Knowing: Toward an Ecological Psychology. Erlbaum; Hillsdale, NJ: 1977. p. 67-82.

Goldman-Rakic PS. The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. Phil Trans R Soc Lond B 1996;351:1445–1453. [PubMed: 8941956]

Goodale MA, Milner AD. Separate visual pathways for perception and action. Trends Neurosci 1992;15:20–5. [PubMed: 1374953]

Griffiths TD, Warren JD. The planum temporale as a computational hub. Trends Neurosci 2002;25:348–53. [PubMed: 12079762]

Griffiths TD, Warren JD. What is an auditory object? Nat Rev Neurosci 2004;5:887–92. [PubMed: 15496866]

Griffiths TD, Rees A, Witton C, Shakir RA, Henning GB, Green GG. Evidence for a sound movement area in the human cerebral cortex. Nature 1996;383:425–7. [PubMed: 8837772]

Griffiths TD, Rees A, Witton C, Cross PM, Shakir RA, Green GG. Spatial and temporal auditory processing deficits following right hemisphere infarction. A psychophysical study. Brain 1997;120:785–94. [PubMed: 9183249]

Griffiths TD, Rees G, Rees A, Green GG, Witton C, Rowe D, Buchel C, Turner R, Frackowiak RS. Right parietal cortex is involved in the perception of sound movement in humans. Nat Neurosci 1998;1:74–9. [PubMed: 10195113]

Grush R. The emulation theory of representation: motor control, imagery, and perception. Behav Brain Sci 2004;27:377–96. discussion 396–442. [PubMed: 15736871]

Guenther FH. Cortical interactions underlying the production of speech sounds. J Commun Disord 2006;39:350–65. [PubMed: 16887139]

Hackett TA. Information flow in the auditory cortical network. Hear Res. 2010

Hackett TA, Stepniewska I, Kaas JH. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 1998a;394:475–95. [PubMed: 9590556]

Hackett TA, Stepniewska I, Kaas JH. Thalamocortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 1998b;400:271–86. [PubMed: 9766404]

Halpern AR, Zatorre RJ. When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. Cereb Cortex 1999;9:697–704. [PubMed: 10554992]

Heffner H, Masterton B. Contribution of auditory cortex to sound localization in the monkey (Macaca mulatta). J Neurophysiol 1975;38:1340–1358. [PubMed: 815514]

Hershberger W. Afference copy, the closed-loop analogue of von Holst's efference copy. Cybernetics Forum 1976;8:97–102.

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. The cortical organization of speech processing. Nature reviews. Neuroscience 2007;8:393–402. [PubMed: 17431404]

Hickok G, Okada K, Serences JT. Area Spt in the human planum temporale supports sensory-motor integration for speech processing. J Neurophysiol 2009;101:2725–32. [PubMed: 19225172]

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. Neurosci Lett 2000;287:156–60. [PubMed: 10854735]

Hikosaka O, Nakahara H, Rand MK, Sakai K, Lu X, Nakamura K, Miyachi S, Doya K. Parallel neural networks for learning sequential procedures. Trends Neurosci 1999;22:464–71. [PubMed: 10481194]

Houde JF, Nagarajan SS, Sekihara K, Merzenich MM. Modulation of the auditory cortex during speech: an MEG study. J Cogn Neurosci 2002;14:1125–38. [PubMed: 12495520]

Humphries C, Liebenthal E, Binder JR. Tonotopic organization of human auditory cortex. Neuroimage 2010;50:1202–11. [PubMed: 20096790]

Hyde KL, Peretz I, Zatorre RJ. Evidence for the role of the right auditory cortex in fine pitch resolution. Neuropsychologia 2008;46:632–9. [PubMed: 17959204]

Indefrey P, Levelt WJM. The spatial and temporal signatures of word production components. Cognition 2004;92:101–144. [PubMed: 15037128]

Irvine, DRF. Physiology of auditory brainstem pathways. In: Fay, RR.; Popper, AA., editors. Springer Handbook of Auditory Research. Volume 2. The Mammalian Auditory Pathway: Neurophysiology. Springer-Verlag; Berlin: 1992. p. 153-231.

Jääskeläinen IP, Ahveninen J, Bonmassar G, Dale AM, Ilmoniemi RJLS, Lin FH, May P, Melcher J, Stufflebeam S, Tiitinen H, Belliveau JW. Human posterior auditory cortex gates novel sounds to consciousness. Proc Natl Acad Sci U S A 2004;101:6809–14. [PubMed: 15096618]

Kaas JH, Hackett TA. Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci U S A 2000;97:11793–9. [PubMed: 11050211]

Kauramäki J, Jääskeläinen IP, Hari R, Möttönen R, Rauschecker JP, Sams M. Transient adaptation of auditory cortex organization by lipreading and own speech production. J Neurosci 2010;30:1314–1321. [PubMed: 20107058]

Kawato M. Internal models for motor control and trajectory planning. Curr Opin Neurobiol 1999;9:718–27. [PubMed: 10607637]

Kayser C, Petkov CI, Augath M, Logothetis NK. Functional imaging reveals visual modulation of specific fields in auditory cortex. J Neurosci 2007;27:1824–35. [PubMed: 17314280]

Keller SS, Roberts N, Hopkins W. A comparative magnetic resonance imaging study of the anatomy, variability, and asymmetry of Broca's area in the human and chimpanzee brain. J Neurosci 2009;29:14607–16. [PubMed: 19923293]

King AJ, Nelken I. Unraveling the principles of auditory cortical processing: can we learn from the visual system? Nat Neurosci 2009;12:698–701. [PubMed: 19471268]

Knudsen EI, Konishi M. Space and frequency are represented separately in auditory midbrain of the owl. J Neurophysiol 1978;41:870–84. [PubMed: 681991]

Krumbholz K, Schönwiesner M, Rübsamen R, Zilles K, Fink GR, von Cramon DY. Hierarchical processing of sound location and motion in the human brainstem and planum temporale. Eur J Neurosci 2005a;21:230–8. [PubMed: 15654860]

Krumbholz K, Schönwiesner M, von Cramon DY, Rübsamen R, Shah NJ, Zilles K, Fink GR. Representation of interaural temporal information from left and right auditory space in the human planum temporale and inferior parietal lobe. Cereb Cortex 2005b;15:317–24. [PubMed: 15297367]

Kusmierek P, Rauschecker JP. Functional specialization of medial auditory belt cortex in the alert rhesus monkey. J Neurophysiol 2009;102:1606–22. [PubMed: 19571201]

Lahav A, Saltzman E, Schlaug G. Action representation of sound: audiomotor recognition network while listening to newly acquired actions. J Neurosci 2007;27:308–14. [PubMed: 17215391]

Lakatos P, Chen CM, O'Connell MN, Mills A, Schroeder CE. Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron 2007;53:279–92. [PubMed: 17224408]

Leaver A, Rauschecker JP. Cortical Representation of Natural Complex Sounds: Effects of Acoustic Features and Auditory Object Category. J Neurosci 2010;30:7604–7612. [PubMed: 20519535]

Leaver A, Van Lare JE, Zielinski BA, Halpern A, Rauschecker JP. Brain activation during anticipation of sound sequences. J Neurosci 2009;29:2477–2485. [PubMed: 19244522]

Lewis JW, Van Essen DC. Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. J Comp Neurol 2000;428:112–37. [PubMed: 11058227]

Lewis JW, Brefczynski JA, Phinney RE, Janik JJ, DeYoe EA. Distinct cortical pathways for processing tool versus animal sounds. J Neurosci 2005;25:5148–58. [PubMed: 15917455]

Liberman AM, Mattingly IG. The motor theory of speech perception revised. Cognition 1985;21:1–36. [PubMed: 4075760]

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M. Perception of the speech code. Psychol Rev 1967;74:431–461. [PubMed: 4170865]

Maeder PP, Meuli RA, Adriani M, Bellmann A, Fornari E, Thiran JP, Pittet A, Clarke S. Distinct Pathways Involved in Sound Recognition and Localization: A Human fMRI Study. Neuroimage 2001;14:802–16. [PubMed: 11554799]

May BJ. Role of the dorsal cochlear nucleus in the sound localization behavior of cats. Hear Res 2000;148:74–87. [PubMed: 10978826]

Micheyl C, Tian B, Carlyon RP, Rauschecker JP. Perceptual organization of tone sequences in the auditory cortex of awake macaques. Neuron 2005;48:139–48. [PubMed: 16202714]

Miller LM, Recanzone GH. Populations of auditory cortical neurons can accurately encode acoustic space across stimulus intensity. Proc Natl Acad Sci U S A 2009;106:5931–5. [PubMed: 19321750]

Morel A, Garraghty PE, Kaas JH. Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. J Comp Neurol 1993;335:437–59. [PubMed: 7693772]

Müller-Preuss P, Ploog D. Inhibition of auditory cortical neurons during phonation. Brain Res 1981;215:61–76. [PubMed: 7260601]

Mulliken GH, Musallam S, Andersen RA. Forward estimation of movement state in posterior parietal cortex. Proc Natl Acad Sci U S A 2008;105:8170–7. [PubMed: 18499800]

Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA 3rd, Brugge JF. Temporal envelope of time-compressed speech represented in the human auditory cortex. J Neurosci 2009;29:15564–74. [PubMed: 20007480]

Numminen J, Salmelin R, Hari R. Subject's own speech reduces reactivity of the human auditory cortex. Neurosci Lett 1999;265:119–22. [PubMed: 10327183]

Obleser J, Zimmermann J, Van Meter J, Rauschecker JP. Multiple stages of auditory speech perception reflected in event-related FMRI. Cereb Cortex 2007;17:2251–7. [PubMed: 17150986]

Obleser J, Boecker H, Drzezga A, Haslinger B, Hennenlotter A, Roettinger M, Eulitz C, Rauschecker JP. Vowel sound extraction in anterior superior temporal cortex. Hum Brain Mapp 2006;27:562–571. [PubMed: 16281283]

Paus T, Perry DW, Zatorre RJ, Worsley KJ, Evans AC. Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor-to-sensory discharges. Eur J Neurosci 1996;8:2236–46. [PubMed: 8950088]

Perry DW, Zatorre RJ, Petrides M, Alivisatos B, Meyer E, Evans AC. Localization of cerebral activity during simple singing. Neuroreport 1999;10:3979–84. [PubMed: 10716244]

Petrides M, Pandya DN. Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. J Comp Neurol 1984;228:105–16. [PubMed: 6480903]

Petrides M, Pandya DN. Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. PLoS Biol 2009;7:e1000170. [PubMed: 19668354]

Pizzamiglio L, Aprile T, Spitoni G, Pitzalis S, Bates E, D'Amico S, Di Russo F. Separate neural systems for processing action- or non-action-related sounds. Neuroimage 2005;24:852–61. [PubMed: 15652320]

Rauschecker JP. Processing of complex sounds in the auditory cortex of cat, monkey and man. Acta Otolaryngol 1997;532:34–38.

Rauschecker JP. Parallel processing in the auditory cortex of primates. Audiol Neurootol 1998;3:86–103. [PubMed: 9575379]

Rauschecker JP. Neural encoding and retrieval of sound sequences. Ann N Y Acad Sci 2005;1060:125–35. [PubMed: 16597759]

Rauschecker, JP. Cortical processing of auditory space: pathways and plasticity. In: Mast, F.; Jäncke, L., editors. Spatial Processing in Navigation, Imagery and Perception. Springer-Verlag; New York: 2007. p. 389-410.

Rauschecker JP, Tian B. Mechanisms and streams for processing of “what” and “where” in auditory cortex. Proc Natl Acad Sci U S A 2000;97:11800–6. [PubMed: 11050212]

Rauschecker JP, Scott SK. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. Nat Neurosci 2009;12:718–24. [PubMed: 19471271]

Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. Science 1995;268:111–4. [PubMed: 7701330]

Rauschecker JP, Tian B, Pons T, Mishkin M. Serial and parallel processing in rhesus monkey auditory cortex. J Comp Neurol 1997;382:89–103. [PubMed: 9136813]

Ravizza RJ, Masterton B. Contribution of neocortex to sound localization in opossum (Didelphis virginiana). J Neurophysiol 1972;35:344–356. [PubMed: 5029954]

Reale RA, Calvert GA, Thesen T, Jenison RL, Kawasaki H, Oya H, Howard MA, Brugge JF. Auditory-visual processing represented in the human superior temporal gyrus. Neuroscience 2007;145:162–84. [PubMed: 17241747]

Recanzone GH. Spatial processing in the auditory cortex of the macaque monkey. Proc Natl Acad Sci U S A 2000;97:11829–35. [PubMed: 11050216]

Recanzone GH, Guard DC, Phan ML, Su TK. Correlation between the activity of single auditory cortical neurons and sound-localization behavior in the macaque monkey. J Neurophysiol 2000;83:2723–39. [PubMed: 10805672]

Remedios R, Logothetis NK, Kayser C. An auditory region in the primate insular cortex responding preferentially to vocal communication sounds. J Neurosci 2009;29:1034–45. [PubMed: 19176812]

Repp BH. Sensorimotor synchronization: a review of the tapping literature. Psychonomic Bulletin & Review 2005;12:969–92. [PubMed: 16615317]

Rilling JK, Glasser MF, Preuss TM, Ma X, Zhao T, Hu X, Behrens TE. The evolution of the arcuate fasciculus revealed with comparative DTI. Nat Neurosci 2008;11:426–8. [PubMed: 18344993]

Rizzolatti G, Ferrari PF, Rozzi S, Fogassi L. The inferior parietal lobule: where action becomes perception. Novartis Found Symp 2006;270:129–40. discussion 140-5, 164–9. [PubMed: 16649712]

Romanski LM, Averbeck BB. The primate cortical auditory system and neural representation of conspecific vocalizations. Annu Rev Neurosci 2009;32:315–46. [PubMed: 19400713]

Romanski LM, Tian B, Fritz J, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 1999;2:1131–6. [PubMed: 10570492]

Sabes PN. The planning and control of reaching movements. Curr Opin Neurobiol 2000;10:740–6. [PubMed: 11240283]

Saur D, Kreher BW, Schnell S, Kummerer D, Kellmeyer P, Vry MS, Umarova R, Musso M, Glauche V, Abel S, Huber W, Rijntjes M, Hennig J, Weiller C. Ventral and dorsal pathways for language. Proc Natl Acad Sci U S A 2008;105:18035–40. [PubMed: 19004769]

Schubotz RI, Friederici AD, von Cramon DY. Time perception and motor timing: a common cortical and subcortical basis revealed by fMRI. Neuroimage 2000;11:1–12. [PubMed: 10686112]

Schubotz RI, von Cramon DY, Lohmann G. Auditory what, where, and when: a sensory somatotopy in lateral premotor cortex. Neuroimage 2003;20:173–85. [PubMed: 14527579]

Scott SK. Auditory processing--speech, space and auditory objects. Curr Opin Neurobiol 2005;15:197–201. [PubMed: 15831402]

Scott SK, Blank CC, Rosen S, Wise RJ. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 2000;123(Pt 12):2400–6. [PubMed: 11099443]

Seltzer B, Pandya DN. Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. J Comp Neurol 1994;343:445–63. [PubMed: 8027452]

Simon, D. Optimal State Estimation. Wiley; New York: 2006.

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;502:894–923. [PubMed: 17447261]

Smith KR, Hsieh IH, Saberi K, Hickok G. Auditory spatial and object processing in the human planum temporale: no evidence for selectivity. J Cogn Neurosci 2009;22:632–9. [PubMed: 19301992]

Sperry RW. Neural basis of the spontaneous optokinetic response produced by visual inversion. J Comp Physiol Psychol 1950;43:482–9. [PubMed: 14794830]

Spierer L, Bernasconi F, Grivel J. The temporoparietal junction as a part of the “when” pathway. J Neurosci 2009a;29:8630–2. [PubMed: 19587268]

Spierer L, Bellmann-Thiran A, Maeder P, Murray MM, Clarke S. Hemispheric competence for auditory spatial representation. Brain 2009b;132:1953–66. [PubMed: 19477962]

Tata MS, Ward LM. Early phase of spatial mismatch negativity is localized to a posterior “where” auditory pathway. Exp Brain Res 2005a;167:481–486. [PubMed: 16283399]

Tata MS, Ward LM. Spatial attention modulates activity in a posterior “where” auditory pathway. Neuropsychologia 2005b;43:509–16. [PubMed: 15716141]

Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. Functional specialization in rhesus monkey auditory cortex. Science 2001;292:290–3. [PubMed: 11303104]

Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. Neuroimage 2008;39:1429–43. [PubMed: 18035557]

Ungerleider, LG.; Mishkin, M. Two cortical visual systems. In: Ingle, DJ.; Goodale, MA.; Mansfield, RJW., editors. Analysis of Visual Behaviour. MIT Press; Cambridge, MA: 1982. p. 549-586.

Von Holst E, Mittelstaedt H. Das Reafferenzprinzip (Wechselwirkungen zwischen Zentralnervensystem und Peripherie). Die Naturwissenschaften 1950;37:464–476.

Warren JD, Zielinski BA, Green GGR, Rauschecker JP, Griffiths TD. Analysis of sound source motion by the human brain. Neuron 2002;34:1–20. [PubMed: 11931733]

Warren JE, Wise RJ, Warren JD. Sounds do-able: auditory-motor transformations and the posterior temporal plane. Trends Neurosci 2005;28:636–43. [PubMed: 16216346]

Weeks RA, Aziz-Sultan A, Bushara KO, Tian B, Wessinger CM, Dang N, Rauschecker JP, Hallett M. A PET study of human auditory spatial processing. Neurosci Lett 1999;262:155–8. [PubMed: 10218879]

Wernicke, C. Der aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer. Basis Cohn & Weigert; Breslau: 1874.

Wernicke, C. Lehrbuch der Gehirnkrankheiten für Aerzte und Studirende Verlag Theodor Fischer. Kassel; Berlin: 1881.

Wessinger CM, VanMeter J, Tian B, Van Lare J, Pekar J, Rauschecker JP. Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. J Cogn Neurosci 2001;13:1–7. [PubMed: 11224904]

Wilson SM, Saygin AP, Sereno MI, Iacoboni M. Listening to speech activates motor areas involved in speech production. Nat Neurosci 2004;7:701–2. [PubMed: 15184903]

Wise RJ, Scott SK, Blank SC, Mummery CJ, Murphy K, Warburton EA. Separate neural subsystems within ‘Wernicke’s area’. Brain 2001;124:83–95. [PubMed: 11133789]

Wolpert DM, Ghahramani Z, Jordan MI. An internal model for sensorimotor integration. Science 1995;269:1880–2. [PubMed: 7569931]

Wolpert DM, Doya K, Kawato M. A unifying computational framework for motor control and social interaction. Philos Trans R Soc Lond B Biol Sci 2003;358:593–602. [PubMed: 12689384]

Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci 2002a;6:37–46. [PubMed: 11849614]

Zatorre RJ, Bouffard M, Belin P. Sensitivity to auditory object features in human temporal neocortex. J Neurosci 2004;24:3637–42. [PubMed: 15071112]

Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: auditory-motor interactions in music perception and production. Nat Rev Neurosci 2007;8:547–58. [PubMed: 17585307]

Zatorre RJ, Bouffard M, Ahad P, Belin P. Where is ‘where’ in the human auditory cortex? Nat Neurosci 2002b;5:905–9. [PubMed: 12195426]

Zhenochin S, Fritz J, Tian B, Ojima H, Rauschecker JP. Thalamo-cortical projections underlying differential responsiveness in the lateral belt areas of rhesus monkey auditory cortex. Assoc Res Otolaryngol Abstr 1998;21:146.

Zimmer U, Macaluso E. High binaural coherence determines successful sound localization and increased activity in posterior auditory areas. Neuron 2005;47:893–905. [PubMed: 16157283]

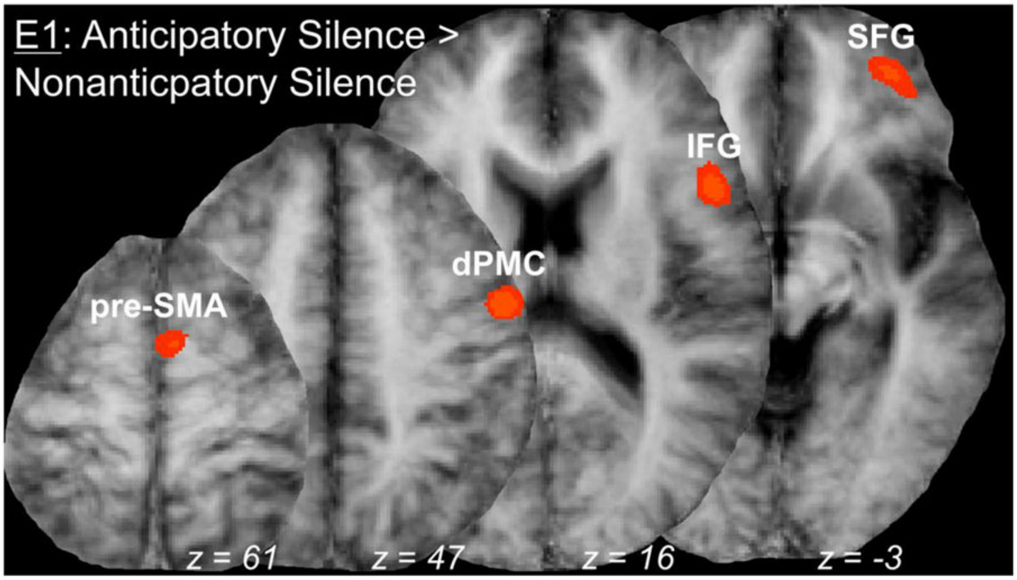

Figure 1. Brain areas active during anticipatory imagery of familiar music

Activated brain regions are found in frontal and premotor regions, including inferior and superior frontal gyrus (IFG, SFG), pre-supplementary motor area (pre-SMA), as well as dorsal and ventral premotor cortex (dPMC, vPMC) (Leaver et al., 2009). Stimuli consisted of the final seconds of familiar or unfamiliar tracks from a compact disk (CD), followed by 8 s of silence. During the silence following familiar tracks from their favorite CD (anticipatory silence, AS, following familiar music, FM), subjects (Ss) reported experiencing anticipatory imagery for each subsequent track. Stimuli presented during unfamiliar trials consisted of music that the Ss had never heard before (unfamiliar music, UM). Thus, during this condition, Ss could not anticipate the onset of the following track (non-anticipatory silence, NS). While in the MRI scanner, subjects were instructed to attend to the stimulus being presented and to imagine, but not vocalize, the subsequent melody where appropriate.

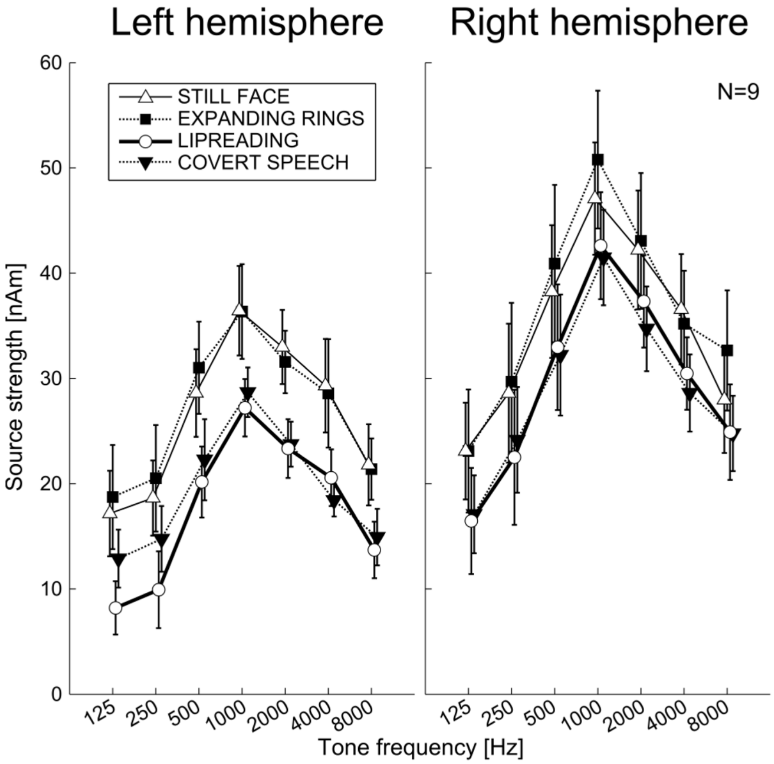

Figure 2. Results of magnetoencephalography (MEG) measuring the effects of lip-reading and covert speech production on human auditory cortex responses

(Kauramäki et al., 2010). Auditory stimuli consisted of 50-ms tones of various frequencies presented in random order. While listening to the tones the subjects (Ss) performed one of four tasks: (1) “lip-reading”, i.e. Ss watched video clips of a face silently articulating Finnish vowels, (2) a visual control task of comparable difficulty (“expanding rings”), (3) a “still-face” passive control condition, and (4) “covert production” of the same vowels. During the still-face and covert-speech conditions, the Ss saw the same static face on the screen. During the expanding-rings as well as lip-reading conditions, Ss performed a one-back task. Auditory-cortex responses with a latency around 100 ms (N100m) were equally suppressed in the lip-reading and covert speech-production tasks compared with the visual control and baseline tasks; the effects involved all frequencies and were most prominent in the left hemisphere. Responses showed significantly increased N100m suppression immediately after the articulatory gesture. These findings suggest that the lip-reading-related suppression in the auditory cortex is caused by an efference copy from the speech-production system, generated during both own speech and lip-reading.

The lower panel shows the mean (± standard error) differences in active task conditions relative to the passive still-face baseline. Asterisks indicate significant differences at a given frequency between the lip-reading *vs.* expanding-rings tasks (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

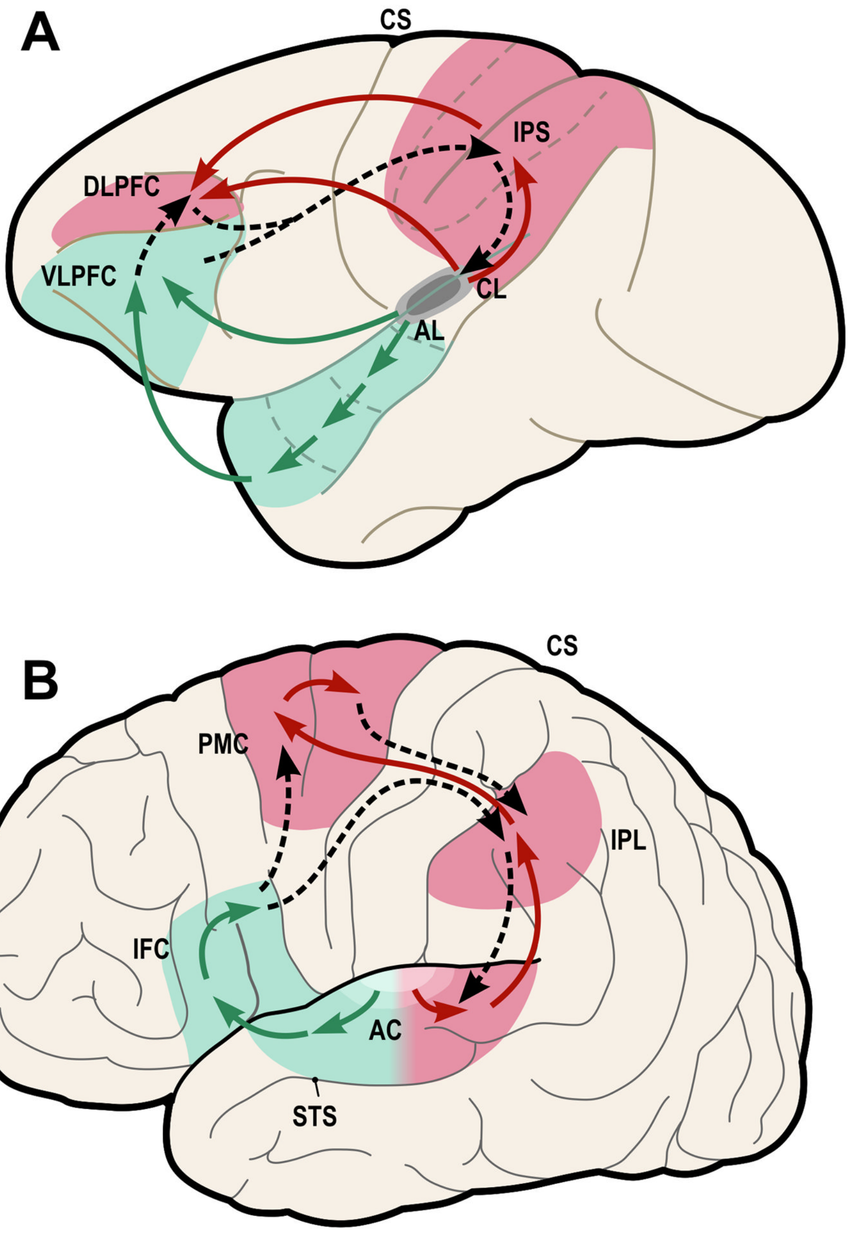

Figure 3. Expanded model of dual auditory processing streams in the primate brain: a) Rhesus monkey (modified from Rauschecker and Tian, 2000); b) Human (simplified from Rauschecker and Scott, 2009)

While the role of the antero-ventral stream (green) in auditory object recognition, including perception of vocalizations and speech, is now widely accepted, the exact role of the postero-dorsal (or just "dorsal") stream (red) is still being debated. Its function clearly includes spatial processing, but a role in human speech and language has also long been postulated. A reinterpretation of these classical studies suggests that the dorsal stream pivots around inferior/posterior parietal cortex, where a quick sketch of sensory event information is compared with an efference copy of motor plans (dashed lines). Thus, the dorsal stream plays a more general role in sensorimotor integration and control.

In clockwise fashion, starting out from auditory cortex, the processing loop performs as a forward model: Object information, such as vocalizations and speech, is decoded in the antero-ventral stream all the way to category-invariant inferior frontal cortex (IFC, or VLPFC in monkeys) and transformed into articulatory representations (DLPFC or ventral PMC). Frontal activations are transmitted to the IPL and pST, where they are compared with auditory and other sensory information. It is this fronto-parietal-sensory section that turns the dorsal stream on its head and expands its fundtion.

AC: auditory cortex; STS: superior temporal sulcus; IFC: inferior frontal cortex; PFC: prefrontal cortex; PMC: premotor cortex; IPL: inferior parietal lobule; IPS: inferior parietal sulcus; CS: central sulcus.