

Chapter 2

Processing Streams in Auditory Cortex

1

2

Josef P. Rauschecker

3

Keywords

Anterior ectosylvian • Bandpass noise • Combination sensitivity • 4

Frequency modulation • Functional MRI • Inferior frontal • Inferior parietal lobule 5

• Lateral belt • Magnetoencephalography • Medial geniculate • Monkey calls • PET 6

imaging • Prefrontal cortex • Premotor cortex • Superior temporal 7

Abbreviations

8

A1	primary auditory cortex	9
AEA	anterior ectosylvian auditory area	10
AES	anterior ectosylvian sulcus	11
AL	anterolateral area	12
aST	anterior superior temporal	13
BA	Brodmann area	14
BOLD	blood oxygen level–dependent	15
BP	band-passed	16
BPN	band-passed noise	17
CD	compact disk	18
CL	caudolateral area	19
CM	caudomedial belt field	20
cs	central sulcus	21
CS	combination sensitivity	22
DCN	dorsal cochlear nucleus	23
DLPFC	dorsolateral prefrontal cortex	24
dPMC, vPMC	dorsal and ventral premotor cortex	25

J.P. Rauschecker (✉)

Department of Neuroscience, Georgetown University Medical Center,

New Research Building, Room WP19, Box 571460, Washington, DC 20057, USA

e-mail: rauschej@georgetown.edu

26	FM	frequency-modulated
27	fMRI	functional magnetic resonance imaging
28	IFC	inferior frontal cortex
29	IFG, SFG	inferior and superior frontal gyrus
30	IPL	inferior parietal lobule
31	ITDILD	interaural time and level differences
32	LB, MB	lateral and medial belt
33	LIP	lateral intraparietal area
34	MCs	monkey calls
35	MCPI	monkey call preference index
36	MEG	magnetoencephalography
37	MGd	dorsal nucleus of the medial geniculate
38	MGM	medial nucleus of the medial geniculate
39	ML	middle lateral area
40	MMN	mismatch negativity
41	MSO, LSO	medial and lateral superior olive
42	NSF	National Science Foundation
43	PAF	posterior auditory field
44	PET	positron emission tomography
45	PFC	prefrontal cortex
46	PMC	premotor cortex
47	PPC	posterior parietal cortex
48	pre-SMA	presupplementary motor area
49	pST	posterior superior temporal
50	PT	planum temporale
51	R, RL	rostral (=rostrolateral) field
52	SC	superior colliculus
53	STG	superior temporal gyrus
54	STP	supratemporal plane

2.1 Hierarchical Processing of Communication Sounds in the Auditory Ventral Stream

2.1.1 Auditory Communication as a Pattern Recognition Problem

Auditory communication involves the decoding of complex sounds and the assignment of specific sounds to behaviorally relevant meanings. While the latter problem has long been discussed in the realms of ethology and linguistics, the former should be solvable more immediately by neurophysiology. But even the phonetic decoding problem is a difficult one, comparable in complexity with visual pattern recognition. Like in vision, it involves the combination of a multitude of features that occur simultaneously and in temporal sequence. The unique combination of features in the spectral and temporal domain, therefore, characterizes a specific communication sound. This alone, however, is not enough. The feature representation has to be

robust and invariant against spurious changes and distortions caused by unpredictable influences.

The way such feature representations are generated in the brain is equally unclear. It could be that neurons in the auditory pathways become more and more specific for frequency, as one ascends the different levels of processing from the cochlea to auditory cortex and beyond. Such highly frequency-specific neurons could then be simultaneously active in concert when a complex sound, such as a communication sound, is present in the animal's environment. The opposite, however, is conceivable as well: Neurons at higher levels of the auditory pathways could become more and more broadly tuned for frequency, as they combine more and more inputs across the frequency domain. As a consequence, they would become more and more specific for certain sounds and unresponsive to others.

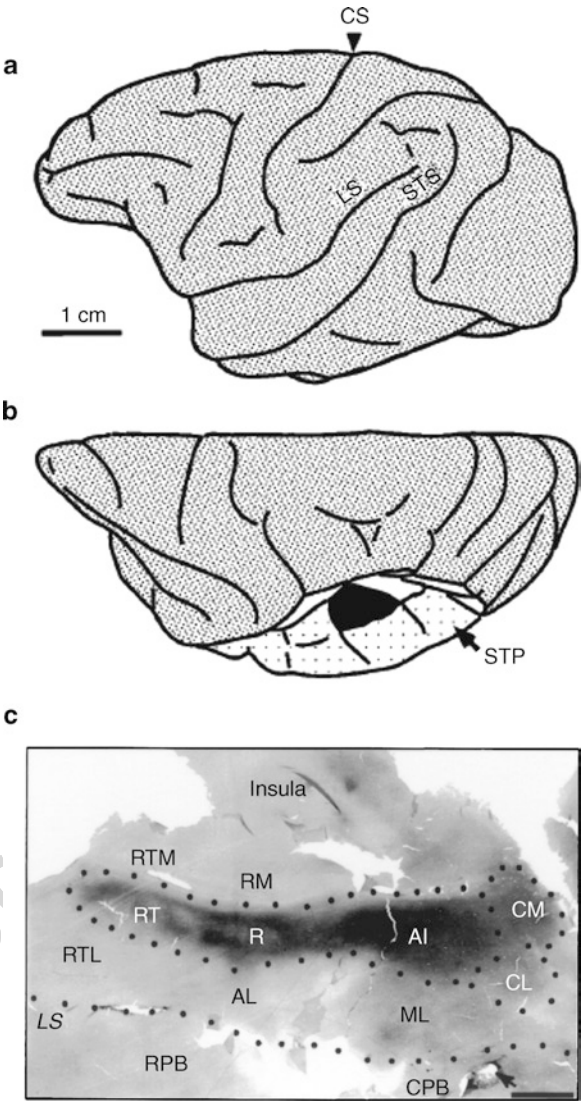
It appears that both processes are in fact happening. While the tuning of neurons for frequency is still relatively broad at the level of the auditory nerve, it becomes sharpened in the lemniscal pathways of brain stem and thalamus, with neurons in primary auditory cortex (A1) very narrowly tuned for frequency. As one travels beyond A1, however, the trend reverses dramatically. One of the discoveries of Merzenich and Brugge in their early mapping studies of auditory cortex in the rhesus monkey (*Macaca mulatta*) (Fig. 2.1a, b) was that neurons in the nonprimary, lateral part of the supratemporal plane (STP) responded only poorly to pure tones and were, therefore, hard to characterize (Brugge & Merzenich, 1973; Merzenich & Brugge, 1973). Merzenich and Brugge's assumption was that the neurons in these lateral fields might respond to more complex sounds. As was discovered later, this is indeed the case (Rauschecker et al., 1995): Neurons in the lateral belt (LB) clearly prefer band-passed (BP) noise bursts to tone pips and also respond well to other types of complex sounds, as reviewed here.

Therefore, the coding problem of auditory pattern recognition does not end at the level of primary auditory cortex; it actually begins there. The realization of the existence of a multitude of auditory cortical representations (Merzenich & Brugge, 1973; Reale & Imig, 1980) on one hand adds to the complexity of the problem but on the other hand may be an important part of the solution. Could the various cortical fields be specialized for certain aspects of the auditory world and represent different levels of analysis? This is the answer suggested by research on visual pattern and object recognition over the last 30 years, in which it has become abundantly clear that different cortical areas do indeed specialize in certain aspects of the visual world. The specialization is, of course, not absolute—in fact, there exists substantial overlap—but there is clear evidence for “families” of specialized areas that are organized into processing streams emphasizing distinct aspects of our natural surroundings.

2.1.2 Early Parallel Processing in Auditory Cortex 104

Parallel processing streams in auditory cortex start as early as the core areas: Area A1 and the rostral auditory area (R) are both koniocortical areas with neurons sharply tuned for frequency and tonotopic maps that are mirror-symmetric. Combined lesion and tracer studies (Rauschecker et al., 1997) have shown that both

Fig. 2.1 Location of auditory cortical areas in the macaque brain. Left hemisphere is displayed. **(b)** Primary auditory cortex is shown marked in black on the supratemporal plane (STP) inside the lateral sulcus (LS). STS, superior temporal sulcus; CS, central sulcus. **(a, b)** adapted from Merzenich & Brugge, 1973). **(c)** Histochemical identification of core and belt areas by parvalbumin staining in a flatmount of rhesus monkey auditory cortex. **(c** from Hackett et al., 1998)



cortical core areas receive input from the principal relay nucleus of the auditory thalamus, the ventral nucleus of the medial geniculate (MGv). By contrast, the other prominent area on the supratemporal plane of the macaque, the caudomedial area (CM), does not receive input from MGv but only from the medial and dorsal subnuclei of the medial geniculate (MGd and MGm). As a consequence, lesions of A1 lead to unresponsiveness of neurons in CM to tonal stimulation, but not of neurons in area R, which receive independent input from MGv. To be sure, the parallel input to areas of the supratemporal plane may start even more peripherally than the thalamus. Studies of the auditory brain stem indicate that the ventral and dorsal cochlear

nuclei (VCN, DCN) have very different response characteristics and may ultimately give rise to different functions of hearing, including auditory pattern and space processing, respectively (Yu & Young, 2000).

2.1.3 Processing of Sounds with Intermediate Complexity in the Auditory Belt

As has been recognized early on the basis of cytoarchitectonics (Pandya & Sanides, 1972), the auditory region in the superior temporal cortex consists of a “core” with a koniocortical appearance surrounded by a “belt.” The cytoarchitecture is matched by distinct histochemical differences (Kaas & Hackett, 2000; Hackett, 2011) that make the core stand out by dark staining compared to the belt (with intermediate staining) and another zone termed “parabelt” (PB) (Morel et al., 1993; Hackett et al., 1998) with very light staining (Fig. 2.1c) (see also Jones et al., 1995). Some processing characteristics of the belt are discussed first and are then compared with those of the core.

2.1.3.1 Selectivity for BP Noise

One fundamental finding that was secured in initial studies (Rauschecker et al., 1995) and confirmed in detail later (Rauschecker & Tian, 2004) was the enhanced response of LB neurons to BP noise compared to pure tones (Fig. 2.2a, b). This demonstrated the ability of LB neurons to integrate over a finite frequency spectrum in a facilitatory fashion. By comparison, this integrative ability is largely absent in A1 neurons, a significant difference that we return to later.

The finding of robust auditory responses to band-passed noise (BPN) stimuli in LB neurons also was of great practical value, however, because it permitted systematic mapping of the LB. BPN bursts have a clearly defined center frequency as well as a defined bandwidth. Mapping of the LB along the rostrocaudal dimension reveals a smooth gradient for best center frequency with two reversals (Rauschecker et al., 1995; Rauschecker & Tian, 2004). This means that there exist three cochleotopically organized areas within the LB, which were termed the anterolateral, middle lateral, and caudolateral areas (AL, ML, and CL; Fig. 2.2c).

Not only do LB neurons integrate over frequency, but they do so in a rather specific way, which produces the best response at a specific “best bandwidth” (BBW) (Fig. 2.3a). Presumably, this is the result of intricate interactions between excitatory and inhibitory inputs. BBWs in the LB are about equally distributed over the whole bandwidth spectrum, whereas A1 neurons clearly prefer pure tones to BP noise. Thus, there was a clear trend for BBW to increase from core toward belt (Rauschecker & Tian, 2004). The same was later found for medial belt (Kusmierek & Rauschecker, 2009).

Neurons with selectivity for the center frequency and bandwidth of BPN bursts are ideally suited to participate in the decoding of communication sounds. Such sounds contain many instances of BPN bursts (Fig. 2.4), not only in rhesus monkeys, but also in many other species (Wang, 2000), including humans. BPN detectors

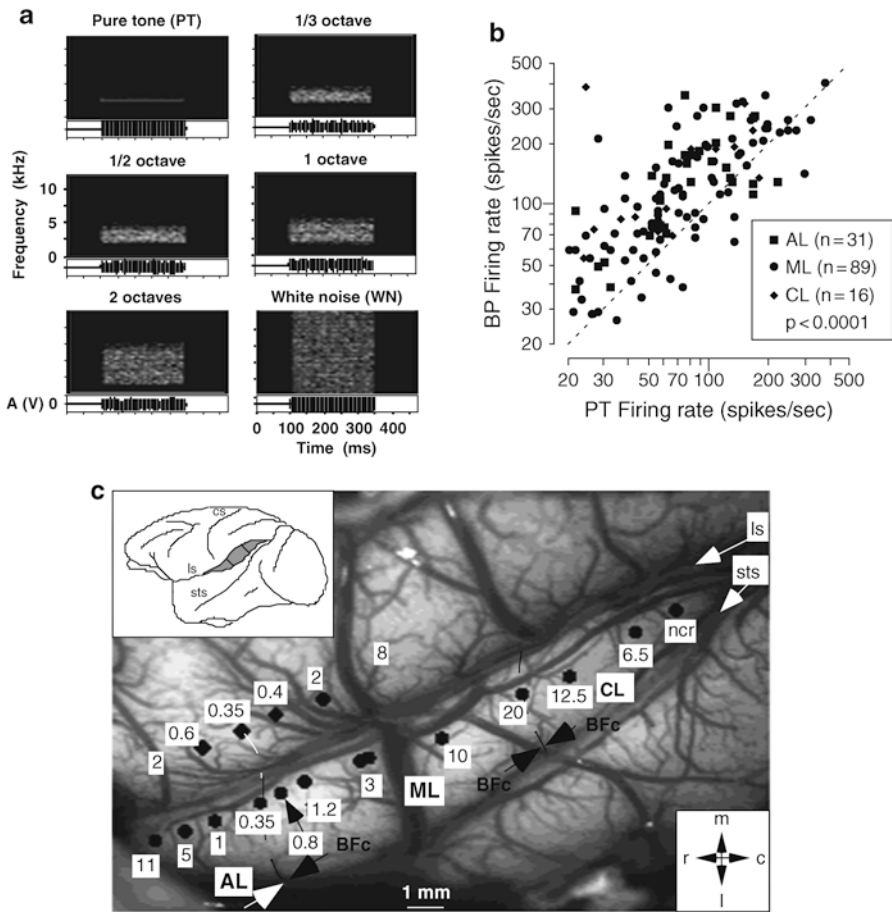


Fig. 2.2 Characterization of lateral belt (LB) areas in rhesus monkey auditory cortex. (**a**, **b**) Preference of neurons in the three LB areas (AL, ML, and CL) for band-passed (BP) noise bursts over tone pips (PT). Almost invariably BP stimuli elicit a higher spike rate in the same neurons than PT stimuli. (**c**) Cochleotopic organization of LB areas. Is, lateral sulcus; cs, central sulcus; sts, superior temporal sulcus. (From Rauschecker et al., 1995)

Fig. 2.3 (continued) rate after subtracting spontaneous (base) firing rate (shown in open symbols). (From Rauschecker & Tian, 2004.) (**b**, **c**) Response of a typical LB neuron to FM sweeps of different rate and direction. Peristimulus time histograms and raster dot displays are shown above a schematic display of the respective stimuli. Upward FM directions are shown in (**b**), downward directions in (**c**). FM rates are displayed on top of each column. A clear preference for a highly specific FM rate and for FM sweeps in the upward direction is found. (From Tian & Rauschecker, 2004)

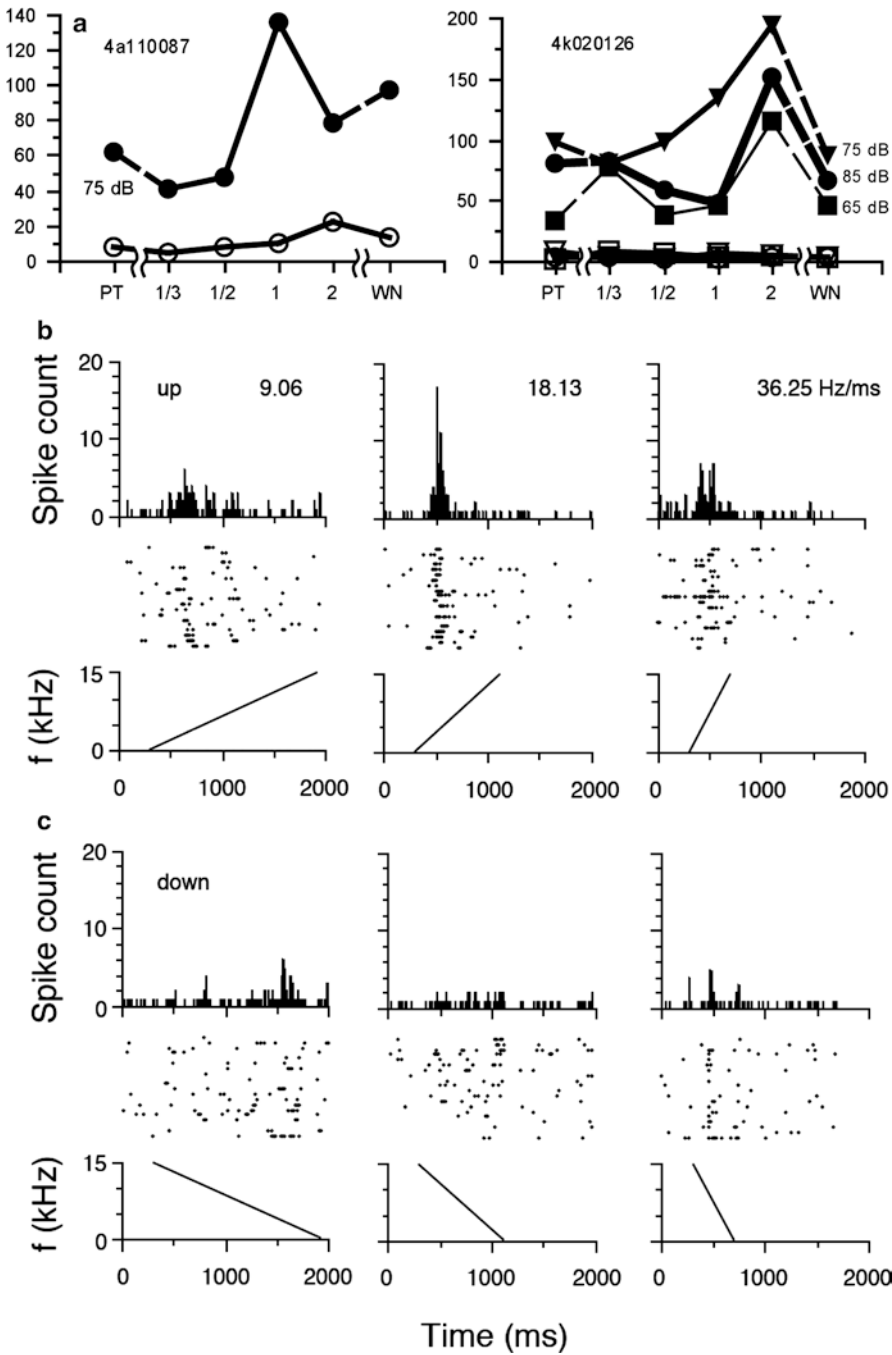


Fig. 2.3 Selectivity of LB neurons for sounds of intermediate complexity. **(a)** Responses (in spikes per second) of two auditory cortical neurons to BP noise bursts with different bandwidth. Maximum firing rate is plotted diagrammatically against bandwidth. Filled symbols show net firing

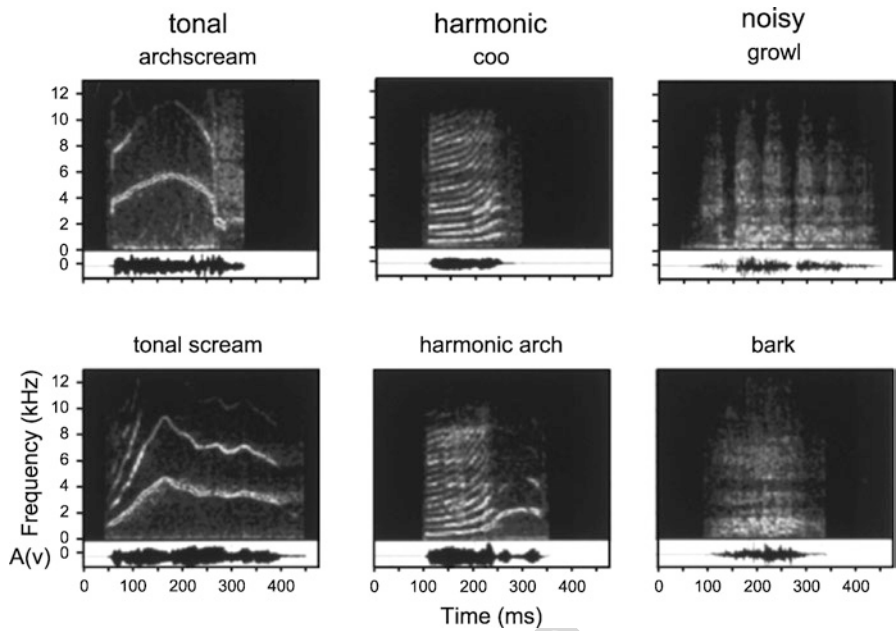


Fig. 2.4 Six examples of rhesus monkey calls in three phonetic-acoustic categories. First column: tonal calls; second column: harmonic calls; third column: noisy calls. Spectrograms are shown with corresponding time signals underneath. (From Rauschecker, 1998b)

would, therefore, almost have to be included in the repertoire of feature detectors dealing with communication sounds. To perform such a task adequately, however, the feature detectors would have to preserve their selectivity regardless of sound intensity. Indeed, as one of the examples in Fig. 2.3a shows, LB neurons generally do prefer the same center frequency at different intensities.

2.1.3.2 Selectivity for Frequency-Modulated Sweeps

Other features that are highly typical for communication sounds in most species are changes in frequency over time (“frequency-modulated [FM] sweeps”), also sometimes referred to as chirps or glides. FM sweeps are characterized by two parameters: FM rate and direction. Neurons in the LB are highly selective for both parameters (Tian & Rauschecker, 2004). First, 94% of LB neurons responded to FM stimuli in at least one direction. To characterize FM direction selectivity (DS) quantitatively, a DS index was calculated. A neuron was considered direction selective when the response in one FM direction for one or more FM rate was at least twice as large as that in the other direction (Mendelson & Cynader, 1985). An example is shown in Fig. 2.3b and c. About 60% of LB neurons were classified as direction-selective on the basis of this criterion, with roughly equal proportions of neurons preferring upward and downward directions (Tian & Rauschecker, 2004).

Even more striking was the selectivity of LB neurons for FM rate. Various types of FM-rate tuning can be discerned in the LB, including high-pass, low-pass, and band-pass tuning (Tian & Rauschecker, 2004). Neurons tuned to both FM direction and FM rate, like the one in Fig. 2.3, would be ideal candidates for the extraction of communication-sound features, such as formant transitions in human speech. Preferred FM rate differed markedly between the three LB areas. AL neurons preferred lower FM rates: More than half of AL neurons had their preferred FM rates below 64 Hz/ms, with medians of 25 and 50 Hz/ms for upward and downward sweeps, respectively. CL neurons, in contrast, preferred higher FM rates: About 70% of CL neurons preferred FM rates above 64 Hz/ms, with medians of 160 Hz/ms for both directions. ML neurons preferred FM rates in between.

According to these differences, AL neurons would be very well suited to participate in the decoding of species-specific vocalizations, which range mostly between 8 and 50 Hz/ms (Hauser, 1996; Rauschecker, 1998b). The various harmonics in the widely occurring “coo” calls fall between 10 and 40 Hz/ms. Only some of the “screams” contain FM rates above 100 Hz/ms (tonal scream: 103 Hz/ms; arch scream: 314 and 826 Hz/ms for the downward portion). Some of the neurons in AL do include responses to these faster sweeps. It is noteworthy that screams play an important role as alarm calls, which have to be well localizable by members of the same species.

In its role, area AL can be likened to visual area V4, which contains neurons selective for the size of visual objects (Desimone & Schein, 1987) and plays a pivotal role in the ventral visual “what” stream. Just as inferotemporal cortex, which receives input from V4, constitutes the later stages of visual object recognition (Lueschow et al., 1994), neurons in the rostral parabelt (and further anterior in the STG) are expected to rely on input from AL, compute invariances against distortions, and ensure perceptual constancy.

2.1.4 Selectivity for Species-Specific Calls

Neurons in the LB responded more vigorously to time-variant FM sweeps than to tones of constant frequency, and FM sweeps were also generally more effective than BPN bursts. LB neurons were also tested directly with whole monkey calls (MC; Fig. 2.4) or components thereof (Fig. 2.5). Just as BPN bursts and FM sweeps, MC stimuli elicited more vigorous responses in LB than pure tones. MC stimuli were also generally more effective than BPN bursts but not necessarily more so than FM sweeps, which often remained the best stimuli.

2.1.4.1 Nonlinear Integration Mechanisms

LB neurons responded differentially to different types of MC. Although calls often had the same or comparable bandwidths, neuronal responses differed. Response selectivity, therefore, must be based on features contained in the phonetic fine

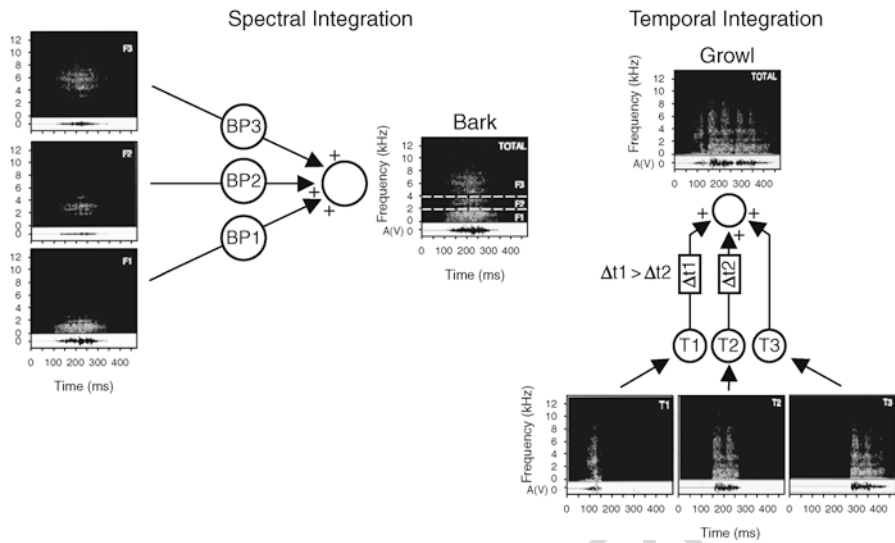


Fig. 2.5 Models of nonlinear spectral and temporal integration (“combination sensitivity”) in neurons of the LB (and presumably PB). Selectivity for complex sounds, such as specific communication calls, is created by combining inputs from lower-order neurons that are BP-selective and/or FM-rate and FM-direction-selective. Delay lines need to be implemented, as explained in the text, to create temporal combination sensitivity. (Extracted from Rauschecker & Scott, 2009)

structure of the calls. The conclusion is not far-fetched that it is the combination of features that causes a cell to respond to a specific type of call and not to others. Indeed, two fundamental mechanisms were identified as causing neuronal selectivity: nonlinear summation (a) in the spectral domain (“spectral facilitation,” SFA) and (b) in the temporal domain (“temporal facilitation,” TFA). This corresponds to spectral and temporal combination sensitivity (CS), respectively, as it has been described previously in other species, such as bats, frogs, and songbirds (Suga et al., 1978; Narins & Capranica, 1980; Margoliash & Fortune, 1992).

In spectral CS or SFA, inputs from lower-order neurons, such as band-pass-selective neurons, are combined in the frequency domain (Fig. 2.5, top). In temporal CS or TFA, inputs are combined in the time domain (Fig. 2.5, bottom). However, both mechanisms are based on the same principle: coincidence detection by cell membranes with a relatively high threshold, that is, a logical AND-gate principle. Only with all inputs present simultaneously a response is evoked; with one input alone, no response follows. This explains why single components or syllables within a call usually are not sufficient to elicit a response. Temporal summation is accomplished by introducing staggered delays in the input pathways transmitting the early components, so all inputs eventually arrive simultaneously at the higher-order target neuron.

In some cases, however, the opposite is found: LB neurons respond decently to single components, but the response is suppressed by presenting the whole call. This is referred to as spectral or temporal suppression (SSU or TSU), respectively. We currently do not fully understand the significance of such units but assume that they are part of the logical alphabet implemented at that level.

One of the most striking differences between core and belt areas identified so far is the difference in their ability to nonlinearly integrate information both in the spectral and the time domain. Whereas more than half of the neurons in LB show some form of nonlinear interaction (SFA, SSU; TFA, TSU), only approximately 10% (or less) of the neurons in A1 or R display the same form of behavior. This demonstrates a quantum leap in the processing characteristics of auditory cortex and is one of the strongest arguments for a hierarchical organization in auditory cortex.

2.1.4.2 MC and Spatial Selectivity243

To quantify MC selectivity in different LB areas, a monkey call preference index (MCPI) was calculated depending on the number of calls to which the neuron responds (Tian et al., 2001). In most cases, a standard battery of seven of the most frequently occurring calls was used. An MCPI of seven, therefore, means that the cell responded to all the calls presented. An MCPI of three or less corresponds to a cell that responded to fewer than half of the calls and can be termed “MC-selective,” whereas cells that responded to five or more of the calls are termed “MC-nonselective.” The LB areas differed in their degree of MC selectivity, as quantified on this basis (Tian et al., 2001). Area AL had the greatest percentage of highly selective neurons (MCPI ≤ 2), followed by ML, whereas CL had the smallest percentage of highly selective neurons. Naturally, for the most nonselective neurons (MCPI ≥ 6), the opposite was found: CL had the greatest percentage of such nonselective neurons, AL the least, with ML somewhere between those two extremes. Comparison of MC selectivity showed AL to be more selective than both ML and CL ($p = 0.0006$ and $p = 0.0287$, respectively, Mann–Whitney U -test). This difference was also highly significant when all three areas were compared together ($p = 0.0026$, Kruskal–Wallis, $df = 2$).

Together with the finding that spatial tuning in neurons of the LB shows the opposite areal distribution—highest selectivity is found in CL and lowest in AL (see Section 2.2)—this has led to the hypothesis that AL and CL, which lie on opposite ends of the LB along its rostrocaudal extent, form the beginning of two pathways for the processing of auditory object and space information (Rauschecker & Tian, 2000; Tian et al., 2001). Selectivity along the anteroventral stream increases further toward more anterior locations (Kikuchi et al., 2010). This trend extends all the way to the temporal pole, which is auditorily activated in the macaque and shows a hemispheric difference for species-specific communication sounds (Poremba et al., 2003, 2004).

2.1.5 Auditory Belt Projections to Prefrontal Cortex270

An anatomical study in rhesus monkeys, combined with physiological mapping of LB areas, has demonstrated the existence of largely separate pathways originating in the LB and projecting to different target regions in the prefrontal cortex (Romanski

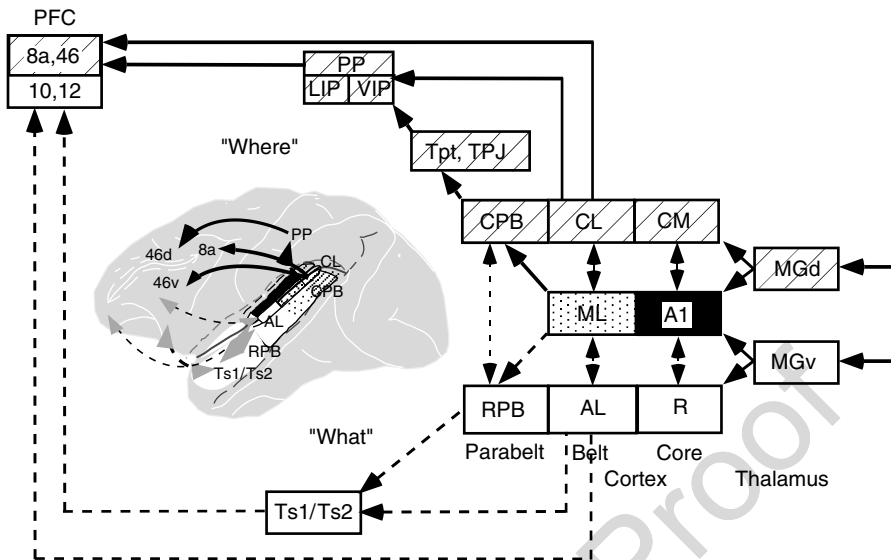


Fig. 2.6 Schematic diagram of dual auditory cortical pathways in primates representing auditory object/pattern (“what”) processing in an anteroventral projection and auditory space (“where”) processing in a posterodorsal projection. (Modified and expanded from Rauschecker, 1998a; Rauschecker & Tian, 2000; Rauschecker, 2007). The projections of the posterodorsal stream are highlighted in solid lines; participating cortical areas are marked with oblique lines. The anteroventral pathway is shown in dashed lines. Areas that are not uniquely participating in either pathway are shown in dark blocks (primary auditory cortex, A1) or stippled (middle lateral belt area, ML). Prefrontal connections of the LB are also shown directly on a lateral view of a rhesus monkey brain. (From Romanski et al., 1999.) MGd, medial geniculate nucleus, dorsal division; MGv, medial geniculate nucleus, ventral division; CM, caudomedial area; R, rostral area; CL, caudolateral area; CPB, caudal parabelt area; RPB, rostral parabelt area; Tpt, temporoparietal junction; PP, posterior parietal cortex; LIP, lateral intraparietal area; VIP, ventral intraparietal area; Ts1, Ts2, rostral temporal areas of Pandya and Sanides (1972); PFC, prefrontal cortex. Brodmann areas are abbreviated with their respective numbers

et al., 1999) (see Fig. 2.6). In this study, three different fluorescent tracers were injected into matched frequency regions of the three LB areas after these had been physiologically mapped. Injections into area AL produced label in ventrolateral and orbital regions of prefrontal cortex, whereas CL injections led to labeling of dorsolateral prefrontal cortex (DLPFC). The latter is known for its involvement in spatial working memory, whereas the former regions are assumed to participate in object working memory (Goldman-Rakic, 1996).

These projection patterns conform to the physiological response properties found in the aforementioned study of Tian et al. (2001), which assigned superior selectivity for auditory patterns and space to areas AL and CL, respectively. The studies by Tian et al. (2001) and Romanski et al. (1999), therefore, form the cornerstones of the theory according to which dual processing streams in nonprimary auditory cortex underlie the perception of auditory objects and auditory space (Rauschecker & Tian,

2000): According to the tracer results, the anteroventral pathway originates from area 287
AL of the LB. Recent physiological data indicate, however, that this pathway may 288
have its origin already in the rostral auditory core area R (Kusmirek et al., 2012). 289
The anteroventral stream projects further from AL via the rostral STG and STS into 290
ventrolateral prefrontal cortex (VLPFC). The anteroventral pathway forms the main 291
substrate for auditory pattern recognition and object identification. An auditory 292
domain is found in VLPFC, in which neurons show responses to complex, nonspa- 293
tial sounds, including animal and human vocalizations (Romanski & Goldman- 294
Rakic, 2002; Cohen et al., 2009). By contrast, another pathway projecting 295
caudodorsally into posterior parietal cortex and DLPFC is thought to be involved in 296
auditory spatial processing (as discussed in Section 2.2). 297

2.1.6 Human Imaging Studies 298

Human neuroimaging studies have confirmed the organization of auditory cortex 299
into core and belt areas by using the same types of stimuli as in the monkey studies 300
(Wessinger et al., 2001; Chevillet et al., 2011). A core region, robustly activated by 301
pure-tone stimuli, is found along Heschl's gyrus. This pure-tone responsive region 302
in human auditory cortex is surrounded by belt regions both medially and laterally, 303
which are activated preferentially by BPN bursts, corresponding to results in nonhu- 304
man primates (Rauschecker et al., 1995; Kusmirek & Rauschecker, 2009). Finally, 305
a region activated by vowel sounds was identified more anterolaterally (Chevillet 306
et al., 2011) (Fig. 2.7). 307

Various findings from human neuroimaging strongly support the dual-stream 308
hypothesis of auditory processing: Anterolateral areas of the superior temporal 309
cortex are activated by intelligible speech (Scott et al., 2000; Binder et al., 2004) or 310
speech-like sounds (Binder et al., 2000; Alain et al., 2001; Maeder et al., 2001), 311
whereas caudal belt and parabelt areas (projecting up dorsally into posterior parietal 312
cortex) are activated by auditory spatial discrimination tasks (see Section 2.2). Some 313
of the areas in anterior human STG do seem to represent species-specific sounds, 314
because they light up only with speech but not with animal calls (Fecteau et al., 315
2004), whereas others may encode more general auditory object information 316
(Zatorre et al., 2004; Leaver & Rauschecker, 2010). 317

Thus it becomes increasingly clear that behaviorally relevant auditory patterns 318
are discriminated in an anterior auditory “what”-stream. Especially with regard to 319
speech it had long been assumed that these processes are located posteriorly in a 320
region called the planum temporale or “Wernicke's area.” These views were largely 321
based on human stroke studies performed more than a century ago (Galaburda et al., 322
1978). By contrast, a recent meta-analysis of human speech processing reviewing 323
more than 100 neuroimaging studies of phoneme, word, and phrase recognition 324
confirms an anterior rather than posterior location of “Wernicke's area”, as defined 325
in this way (DeWitt & Rauschecker, 2012). 326

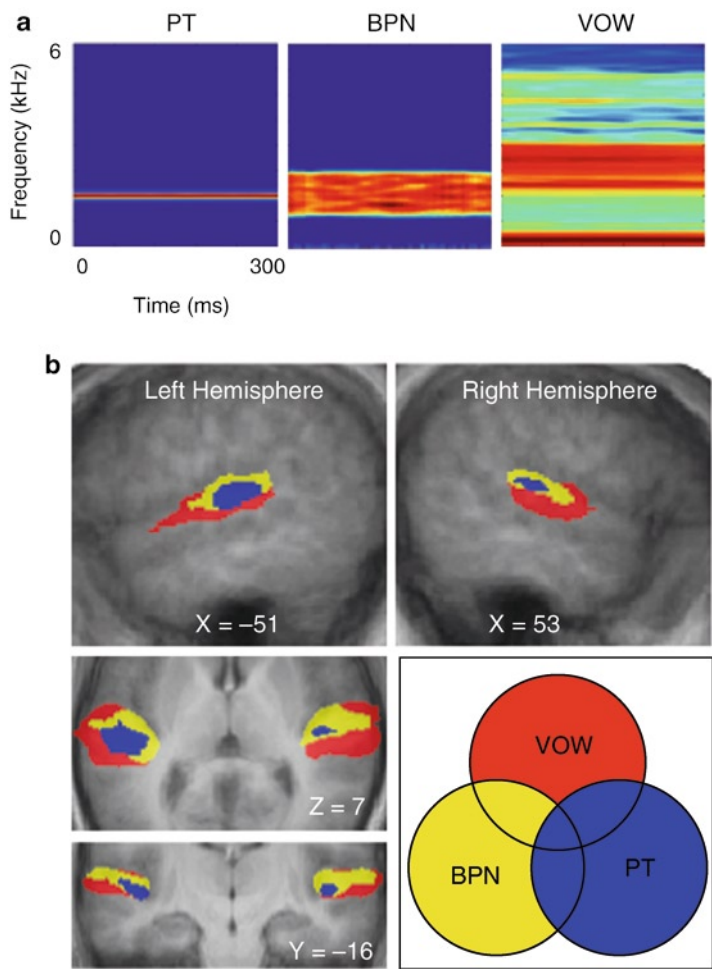


Fig. 2.7 Three functionally discrete regions identified in human auditory cortex using blood oxygen level-dependent (BOLD) imaging and sounds with different spectral complexity. **(a)** Example stimuli used. Subjects were presented with pure tones (PT), band-passed noises (BPN), and species-specific vocalizations (vowel-like speech sounds, VOW), based on the ability to differentiate core, belt and parabelt using these stimulus classes in previous nonhuman primate studies. **(a)** Random-effects group analysis ($n = 13$). A hierarchy of three functionally separable regions is clearly visible in each hemisphere. (From Chevillet et al., 2011)

327 In conclusion, it appears that, like in the visual system, studies of nonhuman
328 primates can serve as excellent models for human studies. Conversely, imaging
329 studies (in both humans and monkeys) can provide useful guidance for microelec-
330 trode studies in nonhuman primates, which permit analyses at much higher spatial
331 and temporal resolution than would be possible in most human studies, with some
332 exceptions (Howard et al., 2000).

2.1.7 Summary 333

Contrary to common belief, which places speech perception in posterior regions of the superior temporal gyrus (STG), converging evidence from three lines of research suggests that communication sounds in both human and nonhuman primates are processed along an anteroventral axis in the STG: (1) Neurophysiological single-unit studies in the rhesus monkey (*Macaca mulatta*) find increased selectivity for species-specific vocalizations along the auditory ventral stream starting with the rostral area (R). (2) The anterolateral area (AL) of auditory belt cortex sends direct anatomical projections to the ventrolateral prefrontal cortex (VLPFC), which has previously been implicated in working memory for patterns and objects. (3) Neuroimaging studies in humans demonstrate that activation associated with the processing of phonemes and words (as well as other auditory “objects”) is consistently localized in areas of the ventral stream.

Thus emerges a picture of an anteriorly directed hierarchical processing stream dedicated to the identification and recognition of behaviorally relevant auditory patterns, which include those used for communication. Activity of neurons in this pathway signals “what” a complex sound represents. Areas R and AL constitute an early stage in this process. Neurons at the level of the belt participate in the decomposition and recombination of auditory features, as they are contained in complex sounds, including conspecific communication calls and human speech. Examples of such features are BP noise bursts and FM sweeps, for which neurons in AL are selective. Neurons in the rostral parabelt (and further anterior in the ventral stream) are expected to rely on input from anterolateral belt, compute invariances against distortions, and ensure perceptual constancy.

2.2 Processing of Space and Motion in the Auditory Dorsal Stream 357 358

2.2.1 Brain Stem Processing of Auditory Space 359

The superior colliculus (SC) is often considered the seat of auditory spatial perception, because it contains a map of auditory space. The SC receives its input from the inferior colliculus (IC), in particular the “external” nuclei of the IC, which in turn receive their input from the dorsal part of the cochlear nuclei (DCN). The medial superior olive (MSO) and the lateral superior olive (LSO) are responsible for encoding interaural time and level differences (ITD and ILD), respectively. For a more complete review of brainstem mechanisms of auditory space processing see Irvine (1992).

2.2.2 *Spatial Selectivity in Primary and Nonprimary Auditory Cortex*

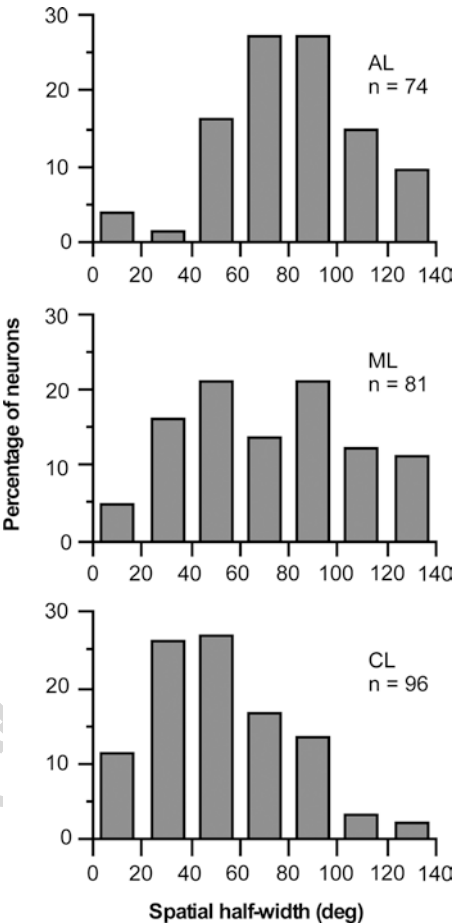
Early studies have suggested a role for auditory cortex in sound localization (Diamond et al., 1956; Ravizza & Masterton, 1972; Heffner & Masterton, 1975). The first study, however, to unequivocally demonstrate that a lesion of primary auditory cortex (A1) in cats causes a deficit in sound localization was performed by Jenkins and Merzenich (1984). Particularly convincing was the fact that the sound localization deficits after small A1 lesions were frequency-specific. These findings were confirmed in later studies using different tasks (Heffner & Heffner, 1990; Beitel & Kaas, 1993). In these studies, A1 appeared to be the only region of auditory cortex whose ablation caused a localization deficit. However, cats have an auditory cortical region that is hidden deep in the anterior ectosylvian sulcus (AES), which was later implicated in sound localization: the anterior ectosylvian auditory area (AEA) (Korte & Rauschecker, 1993; Rauschecker & Korte, 1993; Middlebrooks et al., 1994) or “field AES” (Meredith & Clemo, 1989). It is the main source of auditory cortical input to the SC in cats. In addition, on the basis of cortical cooling studies, the posterior auditory field (PAF) has also been shown to play a role in sound localization by cats (Malhotra et al., 2004; Lomber & Malhotra, 2008).

Spatial tuning of single cortical neurons in cat A1 was measured by a number of groups (Imig et al., 1990; Rajan et al., 1990a, b). They all found two types of spatial tuning: single-peak and hemifield. In single-peak neurons, the best response is found at a particular azimuth location; in hemifield neurons, the response is largely restricted to speaker locations in the contra- or ipsilateral hemifield. In rhesus monkeys, spatially tuned neurons are also found in A1. However, such neurons are found at an even higher rate in the caudal belt (Rauschecker et al., 1997; Recanzone, 2000; Tian et al., 2001). Further, when monkeys are trained in an auditory localization task, the firing rate of caudal belt neurons correlates more tightly with behavioral performance than that of neurons in A1, which is a strong indication that the caudal belt plays an important role in sound localization (Recanzone et al., 2000).

2.2.3 *Spatial and Pattern Selectivity in the Lateral Belt*

In order to compare the spatial selectivity of neurons in the rostral and caudal LB directly in the same animals, broadband species-specific communication calls (MCs) were presented in free field from different locations (Tian et al., 2001) (see Section 2.1). LB areas differed in their degree of MC selectivity, as quantified on the basis of a monkey call preference index (MCPI). The anterolateral area (AL) had the greatest percentage of highly selective neurons ($\text{MCPI} \leq 2$), whereas the caudolateral area (CL) had the smallest percentage of highly MC-selective neurons. Spatial tuning, as measured by the number of speaker locations from which a significant response could be elicited in a free-field experiment, showed the opposite areal distribution in LB than MC: the highest selectivity was found in CL and the lowest in AL (Fig. 2.8).

Fig. 2.8 Distribution of spatial half-width in areas AL, ML, and CL. Summary data from four monkeys are shown in histogram form. Number of units recorded in each area is given on the right. Neurons in CL show significantly greater spatial selectivity than neurons in AL or ML. By contrast, neurons in AL are more selective for monkey calls than neurons in either of the other areas (see Section 2.1). (From Tian et al., 2001)



2.2.4 Caudal Belt Projections to Parietal and Prefrontal Cortex

408

The anatomical tracer study of Romanski et al. (1999) demonstrated the existence of largely segregated pathways from the LB to prefrontal cortex (see Section 2.1 and Fig. 2.6). Injections into area AL produced label in ventrolateral and orbital regions of prefrontal cortex (areas 10, 12), whereas CL injections led to labeling of DLPFC (areas 8a, 46). These projection patterns conformed to the physiological response properties in LB found in the study of Tian et al. (2001), which assigned superior selectivity for auditory patterns and space to areas AL and CL, respectively. The pathway projecting caudodorsally from A1 to the caudal belt (areas CM and CL) and parabelt (areas Tpt and CPB) into posterior parietal cortex (PPC) and DLPFC is thought to be involved in auditory spatial processing.

A projection from posterior STG to PPC in monkeys has been found independently by Lewis and Van Essen (2000). Specifically, the ventral intraparietal area

409
410
411
412
413
414
415
416
417
418
419
420

(VIP) in the PPC has been identified as the primary recipient of auditory input to PPC. The lateral intraparietal area (LIP) has been found to contain auditory neurons as well, but only after training monkeys on auditory saccades (Stricanne et al., 1996; Grunewald et al., 1999; cf. Gifford & Cohen, 2004).

2.2.5 Human Imaging Studies of the Auditory Dorsal Stream

Whereas anterolateral areas of the superior temporal cortex are activated by intelligible speech (see Section 2.1), caudal belt and parabelt areas (projecting up dorsally into PPC) are activated by auditory spatial discrimination tasks (Maeder et al., 2001; Zatorre & Belin, 2001; Arnott et al., 2004) or tasks involving auditory motion (Warren et al., 2002; Krumbholz et al., 2005) (Fig. 2.9).

Auditory activation of the inferior parietal lobule (IPL) has been demonstrated in human imaging studies as well (Bushara et al., 1999; Weeks et al., 1999; Bremmer et al., 2001). By testing the subjects in a visual as well as in an auditory task during the same imaging session, it was shown that IPL does contain a unimodal auditory spatial representation before multisensory convergence occurs in superior parietal cortex (Bushara et al., 1999). A similar conclusion was reached on the basis of clinical and psychophysical studies (Griffiths et al., 1997, 1998).

In a meta-analysis, Arnott et al. (2004) 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. Almost all temporal lobe activity observed during spatial tasks was confined 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. Finally, inferior frontal activity (Brodmann areas [BA] 45 and 47) was reported in only 9% of the spatial studies, but in 56% of the nonspatial studies. These results support an auditory dual-pathway model in humans in which nonspatial sound information (e.g., sound identity) is processed primarily along an anteroventral stream whereas sound location is processed along a posterodorsal stream, that is, within areas posterior to primary auditory cortex.

In a PET study by Zatorre et al. (2002) posterior auditory cortex responded to sounds that varied in their spatial distribution, but only when multiple complex stimuli were presented simultaneously. Consistent with other studies, these authors also found that the right inferior parietal cortex was specifically recruited in localization tasks.

An fMRI study by Krumbholz and co-workers (2005) found that interaural time differences were represented along a posterior pathway comprising the planum temporale (PT) and IPL of the respective contralateral hemisphere. The response

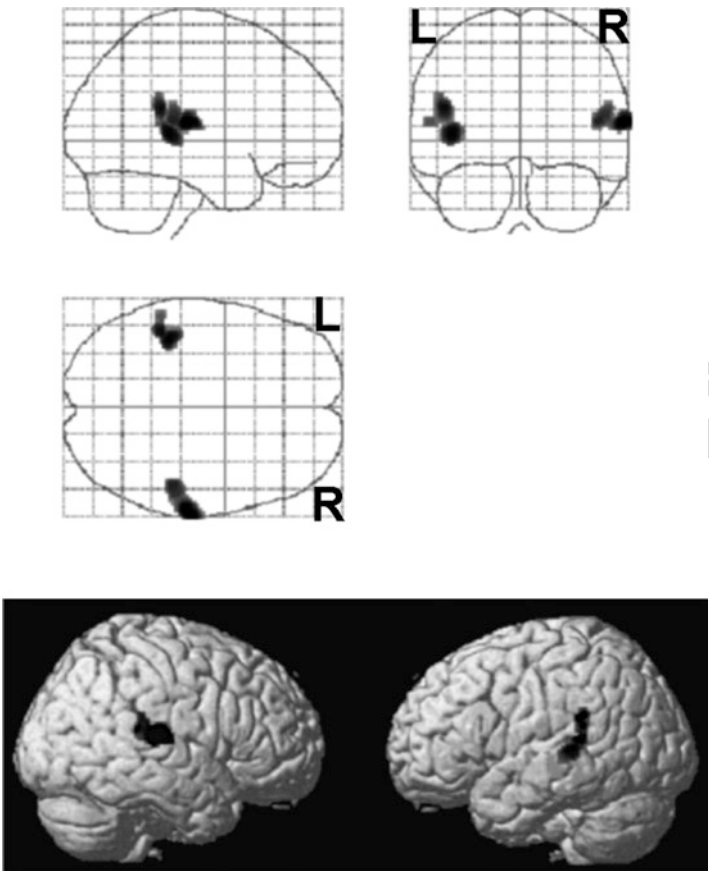


Fig. 2.9 Human imaging of motion in virtual auditory space (all-motion minus stationary-sound). Statistical parametric maps of fMRI group data are shown as sagittal, coronal, and axial projections (above) and rendered onto a canonical brain surface template (below). All voxels significant at the $p < 0.05$ level (corrected for multiple comparisons) are displayed. (From Warren et al., 2002)

was stronger and extended further into adjacent regions of the IPL when the sound was moving than when it was stationary, a finding that confirmed earlier results by Warren et al. (2002) (Fig. 2.9). In contrast to Zatorre et al. (2002), the study by Krumbholz et al. (2005) found that stationary lateralized sounds did produce a significant activation increase in the PT of the respective contralateral hemisphere compared to a centrally presented sound. This discrepancy may be due to the inferior sensitivity of PET relative to fMRI, or to the fact that the spatial ranges of the sounds used by Zatorre et al. (2002) were centered around the midline, and thus always comprised equal parts of both hemifields. This suggests that Zatorre et al. (2002) were unable to detect the contralateral tuning that was observed in the study of Krumbholz et al. (2005).

Timing differences between the two ears can be used to localize sounds in space only when the inputs to the two ears have similar spectrotemporal profiles (high binaural coherence). Zimmer and Macaluso (2005) used fMRI to investigate any modulation of auditory responses by binaural coherence. They assessed how processing of these cues depends on spatial information being task-relevant and whether brain activity correlates with subjects' localization performance. They 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 auditory cortex, but that these cues are used in posterior regions of the STG for successful auditory localization (Zimmer & 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, that is, 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,b) 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 two 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 & Ward, 2005a). Transient attention oriented in cue-target paradigms results in several modulations of the auditory event-related potential. Its earliest component (the Nd1) also reflects modulation of neurons posterior to primary auditory cortex within or near the temporo-parietal junction (TPJ) (Tata & Ward, 2005b).

Selective attention was also used to differentiate the effects of sound location and pitch of an auditory stimulus in an fMRI study (Degerman et al., 2006). Attention to either sound feature produced activation in areas of the superior temporal cortex and in prefrontal and inferior parietal regions. However, during attention to location these activations were located more posterior on the STG than during attention to pitch.

In a study combining fMRI and 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 IPL. Their MEG analysis allowed assessment of the timing of this circuit: Activation of Heschl's gyrus was observed 139 ms after the auditory stimulus, the peak latency of the source located in the posterior STG was at 156 ms, and the IPL and the supramarginal gyrus peaked at 162 ms. Both hemispheres were found to be involved in the processing of sounds coming 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 versus spatial sound changes, demonstrating that the human nonprimary auditory cortex 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.

2.2.6 Summary 525

Contrary to popular belief, which places auditory space processing wholly in the brain stem, several lines of evidence suggest that auditory cortex plays an important role in spatial perception. Lesion studies in animals and humans demonstrate severe deficits in sound localization after damage to auditory cortex. Single-unit recording studies find neurons tuned to spatial location in auditory cortical areas. While these neurons exist already in primary auditory cortex, their prevalence and sharpness of spatial tuning increases in nonprimary areas of the caudal belt, as defined in nonhuman primates. The firing of neurons in the caudal belt also shows a tighter correlation with the behavioral performance of alert monkeys engaged in sound localization behavior. Caudal belt and parabelt project to posterior parietal cortex and to areas of dorsolateral prefrontal cortex, such as the frontal eye and pinna fields, known to be involved in spatial attention and control of eye and pinna movements. This has led to the notion that a posterodorsal processing stream is intimately involved in aspects of auditory spatial processing.

The existence of an auditory “where”-stream is also suggested by functional neuroimaging studies in humans in which subjects process stationary or moving sounds in space. Consistently, posterior aspects of the superior temporal cortex and adjoining inferior parietal areas are activated during these tasks. Thus, although brain stem nuclei perform an important service by computing some of the basic parameters that are prerequisites for spatial processing, such as interaural time and intensity differences, these parameters (together with monaural spectral cues that depend on head and pinnae) are integrated at the cortical level. Auditory space processing, including processing of motion in space, is, therefore, ultimately accomplished at the cortical level.

2.3 Sensorimotor Integration and Control in the Dorsal Stream 550

2.3.1 Introduction 551

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 patterns or “objects,” including recognition of speech sounds; and localization of sounds in space (Rauschecker & Tian, 2000).

Evidence has accumulated, chiefly from work in humans and nonhuman primates, that an anteroventral pathway supports the former function, whereas a posterodorsal stream (including posterior superior temporal cortex, pST, and inferior parietal lobule, IPL) supports the latter, that is, processing of space and motion-in-space. A role of the posterodorsal stream in space and motion would seem to be at odds with classic claims of a function of pST and IPL in speech and language in humans. However, as we have argued previously (Rauschecker & Scott, 2009; Rauschecker, 2011), the neural functions related to space and spoken speech, in a computational sense, may not be as incompatible as they seem. Rather, both functions 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. Although the 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 (Wolpert et al., 1995; Kawato, 1999). In the case of audition, an efference copy sent from prefrontal and premotor cortex provides the basis for “optimal state estimation” in the IPL and in sensory areas of the posterior auditory cortex.

2.3.2 Role of the Human Dorsal Auditory Pathway in Speech and Language

2.3.2.1 Is the Dorsal Pathway Really Involved in Speech Processing?

Section 2.2 of this chapter 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 nonprimate animals). However, another view about the function of pST in humans has classically been even more pervasive: the view that pST is involved in speech or language comprehension (Geschwind, 1965; Damasio & Damasio, 1980). Many textbooks refer to pST and surrounding regions 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 (Wise et al., 2001; Rauschecker & Scott, 2009).

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 (Rilling et al., 2008; Bernal & Ardila, 2009; Keller et al., 2009). 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 & Pandya, 2009), is sparse at best (Frey et al., 2008). Instead, 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 & 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 anteroventral 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.”

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 and IPL are activated by sounds or tasks with other than spatial connotations.

2.3.2.2 Representation of Action Sounds in the Dorsal Stream 617

Various studies have demonstrated activation of left parietal cortical regions while subjects were listening to sounds generated by actions, such as tool sounds (Lewis et al., 2005; Pizzamiglio et al., 2005; Engel et al., 2009). 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 & 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) have also identified activation in the posterior medial PT region within the posterodorsal stream.

Such findings resonate with the “affordance” model of Gibson (1977), in which 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; see also Arbib et al., 2013).

2.3.2.3 A Multisensory Reference Frame 633

The posteromedial region of the PT has been identified as a possible key node for the feedback control of speech production (Dhanjal et al., 2008) because it shows a

response to somatosensory input from articulators as well as to auditory speech input. Adjacent to pST, the temporoparietal 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 caudal belt neurons not only are responsive to auditory input but also reveal multisensory responses (Fu et al., 2003; Brosch et al., 2005; Kayser et al., 2007), although this has been reported about other areas as well (Ghazanfar et al., 2005; Bulkin & Groh, 2006; 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 posterodorsal 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. Some studies suggest that this reference frame transformation begins at earlier stages of the auditory pathway (Mullette-Gillman et al., 2005).

2.3.2.4 Encoding and Retrieval of Sound Sequences

One of the unsolved puzzles in auditory neuroscience is how the brain encodes and stores sequences of sound (Schubotz et al., 2000; Rauschecker, 2005). Unlike tape recorders and compact disk (CD) 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 spectrotemporal 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 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 did not 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

(Wilson et al., 2004; Zatorre et al., 2007; Chen et al., 2008). 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 & 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. 2.10), 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, basal ganglia being 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 (Repp, 2005; Rauschecker & Scott, 2009): 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 determine the exact neurophysiological mechanisms of auditory sequence processing and stream segregation (Micheyl et al., 2005).

2.3.2.5 Auditory Perception/Production Links in Voice and Speech

Monkey studies have shown that neurons in auditory cortex are suppressed during vocalization (Müller-Preuss & Ploog, 1981; Eliades & Wang, 2003). This finding is consistent with results from humans, which indicate that superior temporal areas are suppressed during speech production (Numminen et al., 1999; Curio et al., 2000; Houde et al., 2002). 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).

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 & Scott, 2009). However in nonhuman primate studies, auditory neurons that are suppressed during actual vocalizations are often more activated by distorted vocalizations (Eliades & 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, that is, the angular and supramarginal gyri, in phonology (Caplan et al., 1992), particularly an involvement in the “phonological/articulatory loop” (Baddeley et al., 1984; Aboitiz et al., 2006). This has been confirmed in several functional imaging studies, though the

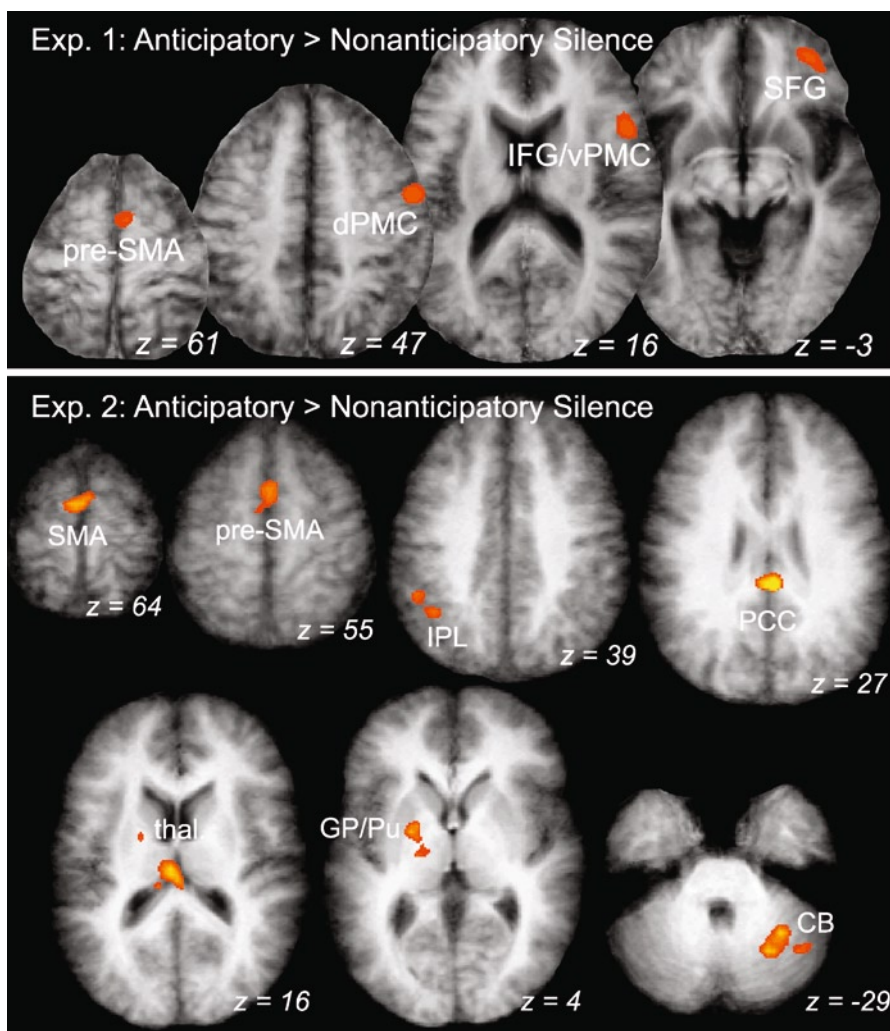


Fig. 2.10 Brain areas active during anticipatory imagery of familiar music. Two experiments were conducted. In Experiment 1 (Exp. 1), stimuli consisted of the final seconds of familiar or unfamiliar tracks from a compact disk (CD), followed by 8 s of silence. During the silence following familiar tracks 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 (nonanticipatory silence, NS). While in the MRI scanner, Ss were instructed to attend to the stimulus being presented and to imagine, but not vocalize, the subsequent melody where appropriate. In Experiment 2 (Exp. 2), Ss were trained in associating newly composed melodies with each other. In both experiments, the outcome was similar: Activated brain regions were found in frontal and premotor regions, including inferior and superior frontal gyrus (IFG, SFG), presupplementary motor area (pre-SMA), as well as dorsal and ventral premotor cortex (dPMC, vPMC). (From Leaver et al., 2009)

precise localization of activity does vary with the type of task used (Gelfand & Bookheimer, 2003; Buchsbaum & D’Esposito, 2008). 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 & Scott, 2009).

New work using diffusion tensor imaging in humans demonstrates that there are direct connections between the pars opercularis of Broca’s area (BA44) and the IPL (Frey et al., 2008; Saur et al., 2008; Bernal & Ardila, 2009), 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 & Pandya, 1984, 2009), and connections between parietal cortex and pST are also well known (Seltzer & Pandya, 1994); together, this could form the basis for a feed-forward network between speech production areas and posterior temporal auditory areas (Fig. 2.11).

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

As documented in Section 2.2, posterior ST regions and the IPL participate in the processing of auditory space and motion. 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 is also activated during listening to music (Lahav et al., 2007; Chen et al., 2008) 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 auditory cortex, which in turn are matched with the corresponding premotor neurons that helped produce the sounds. Thus, specific sensorimotor networks are established that, 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 such as 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 is 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. Although 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

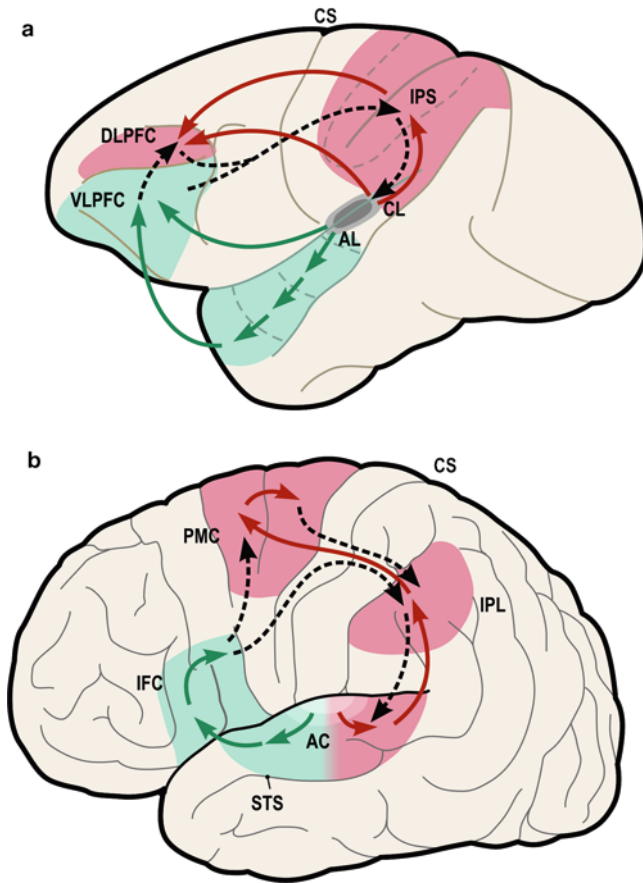


Fig. 2.11 Expanded model of dual auditory processing streams in the primate brain. **(a)** Rhesus monkey. (Modified from Rauschecker & Tian, 2000.) **(b)** Human. (Simplified from Rauschecker & Scott, 2009.) While the role of the anteroventral stream (green) in auditory object recognition, including perception of vocalizations and speech, is now widely accepted, the exact role of the posterodorsal (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 classic 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 anteroventral 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 function. 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

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. 2.11). Learning to produce new sounds in a foreign language should, therefore, lead to changes in both sensory and motor representations of the corresponding sounds.

This basic structure of the underlying fronto-parietal-sensory loops responsible for sensorimotor planning and control is best described as that of an “internal model” or “emulator,” as it is known in motor control theory and robotics (Rauschecker & 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 (Kawato, 1999; Desmurget & Grafton, 2000; Sabes, 2000). More recently, these models have been used to model perception and imagery as well (Wolpert et al., 2003; Grush, 2004). 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.

The feed-forward projection from BA 44 and vPM can be considered the pathway carrying an “efference copy” or “corollary discharge” in the classic sense (Sperry, 1950; Von Holst & 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 visuospatial reaching as well as articulation of speech. At the same time, these sensorimotor loops also support the disambiguation of phonological information.

2.4 Summary

This chapter first summarizes the evidence for the existence of dual pathways in auditory cortex (1) for the processing of pattern/object information and (2) for the processing of space and motion. The former pathway follows an anteroventral route, originating in the rostral field (R) and projecting through rostral belt and parabelt into VLPFC (directly as well as through intermediary stations in the rostral STG). The latter pathway follows a posterodorsal route, originating in primary auditory cortex and projecting through caudal belt and parabelt into DLPFC (directly as well

as via inferior parietal cortex and premotor cortex). Evidence for these pathways comes from anatomical and physiological studies of nonhuman primates as well as neuroimaging studies in humans. A wider consideration of function of the dorsal stream in a third section then discusses the role of the dorsal stream in sensorimotor integration and control more generally. This view incorporates the theoretical concept of internal models.

Future studies of the ventral stream may concentrate on the formation of invariances in the perception of complex sounds, for instance, against changes in pitch. Future work on the dorsal stream should design more specific experiments to test the concept of forward and inverse models and come up with computational formulations of the underlying neural circuitry. This could be done, for instance, by training animals in the learning of auditory sequences, while monitoring neural activity with fMRI as well as chronic microelectrode recordings. Both of these approaches could help to advance our understanding of cognitive aspects of hearing with particular relevance for understanding the evolution/cognition of language and music.

Acknowledgments The present chapter is an updated synthesis of prior publications by the author (Rauschecker & Tian, 2005; Rauschecker, 2007, 2011). 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).

References

- Aboitiz, F., Garcia, R. R., Bosman, C., & Brunetti, E. (2006). Cortical memory mechanisms and language origins. *Brain and Language*, 98(1), 40–56.
- Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., et al. (2006). Task-modulated “what” and “where” pathways in human auditory cortex. *Proceedings of the National Academy of Sciences of the USA*, 103(39), 14608–14613.
- Alain, C., Arnott, S. R., Hevenor, S., Graham, S., & Grady, C. L. (2001). “What” and “where” in the human auditory system. *Proceedings of the National Academy of Sciences of the USA*, 98(21), 12301–12306.
- Arbib, M. A., Verschure, P. F. M. J., & Seifert, U. (2013). Action, Language and Music: Events in Time and Models of the Brain. In M. A. Arbib (Ed.), *Language, music and the brain: A mysterious relationship*. Cambridge, MA: MIT Press.
- Arnott, S. R., Binns, M. A., Grady, C. L., & Alain, C. (2004). Assessing the auditory dual-pathway model in humans. *NeuroImage*, 22(1), 401–408.
- Baddeley, A., Lewis, V., & Vallar, G. (1984). Exploring the articulatory loop. *The Quarterly Journal of Experimental Psychology*, A(36), 233–252.
- Bar, M., Kassam, K. S., Ghuman, A. S., Boshyan, J., Schmid, A. M., Dale, A. M., et al. (2006). Top-down facilitation of visual recognition. *Proceedings of the National Academy of Sciences of the USA*, 103(2), 449–454.
- Beitel, R. E., & Kaas, J. H. (1993). Effects of bilateral and unilateral ablation of auditory cortex in cats on the unconditioned head orienting response to acoustic stimuli. *Journal of Neurophysiology*, 70(1), 351–369.
- Bernal, B., & Ardila, A. (2009). The role of the arcuate fasciculus in conduction aphasia. *Brain*, 132(Pt 9), 2309–2316.

- Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., & Possing, E. T. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cerebral Cortex*, 10(5), 512–528.
- Binder, J. R., Liebenthal, E., Possing, E. T., Medler, D. A., & Ward, B. D. (2004). Neural correlates of sensory and decision processes in auditory object identification. *Nature Neuroscience*, 7(3), 295–301.
- Blakemore, S. J., Goodbody, S. J., & Wolpert, D. M. (1998). Predicting the consequences of our own actions: The role of sensorimotor context estimation. *Journal of Neuroscience*, 18(18), 7511–7518.
- Bremmer, F., Schlack, A., Shah, N. J., Zafiris, O., Kubischik, M., Hoffmann, K., et al. (2001). Polymodal motion processing in posterior parietal and premotor cortex: A human fMRI study strongly implies equivalencies between humans and monkeys. *Neuron*, 29(1), 287–296.
- 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 Anatomique*, 6, 330–357.
- Brosch, M., Selezneva, E., & Scheich, H. (2005). Nonauditory events of a behavioral procedure activate auditory cortex of highly trained monkeys. *Journal of Neuroscience*, 25(29), 6797–6806.
- Brugge, J. F., & Merzenich, M. M. (1973). Responses of neurons in auditory cortex of the macaque monkey to monaural and binaural stimulation. *Journal of Neurophysiology*, 36, 1138–1158.
- Brunetti, M., Belardinelli, P., Caulo, M., Del Gratta, C., Della Penna, S., Ferretti, A., et al. (2005). Human brain activation during passive listening to sounds from different locations: An fMRI and MEG study. *Human Brain Mapping*, 26(4), 251–261.
- Buchsbaum, B. R., & D’Esposito, M. (2008). The search for the phonological store: from loop to convolution. *Journal of Cognitive Neuroscience*, 20(5), 762–778.
- Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds: visual and auditory interactions in the brain. *Current Opinion in Neurobiology*, 16(4), 415–419.
- Bushara, K. O., Weeks, R. A., Ishii, K., Catalan, M.-J., Tian, B., Rauschecker, J. P., & Hallett, M. (1999). Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nature Neuroscience*, 2(8), 759–766.
- Caplan, D., Rochon, E., & Waters, G. S. (1992). Articulatory and phonological determinants of word length effects in span tasks. *The Quarterly Journal of Experimental Psychology*, 45(2), 177–192.
- Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, 18(12), 2844–2854.
- Chevillet, M., Riesenhuber, M., & Rauschecker, J. P. (2011). Functional correlates of the anterolateral processing hierarchy in human auditory cortex. *Journal of Neuroscience*, 31(25), 9345–9352.
- Cohen, Y. E., Russ, B. E., Davis, S. J., Baker, A. E., Ackelson, A. L., & Nitecki, R. (2009). A functional role for the ventrolateral prefrontal cortex in non-spatial auditory cognition. *Proceedings of the National Academy of Sciences of the USA*, 106(47), 20045–20050.
- Curio, G., Neuloh, G., Numminen, J., Jousmaki, V., & Hari, R. (2000). Speaking modifies voice-evoked activity in the human auditory cortex. *Human Brain Mapping*, 9(4), 183–191.
- Cusack, R. (2005). The intraparietal sulcus and perceptual organization. *Journal of Cognitive Neuroscience*, 17(4), 641–651.
- Damasio, H., & Damasio, A. R. (1980). The anatomical basis of conduction aphasia. *Brain*, 103(2), 337–350.
- Davis, B., Christie, J., & Rorden, C. (2009). Temporal order judgments activate temporal parietal junction. *Journal of Neuroscience*, 29(10), 3182–3188.
- Degerman, A., Rinne, T., Salmi, J., Salonen, O., & Alho, K. (2006). Selective attention to sound location or pitch studied with fMRI. *Brain Research*, 1077(1), 123–134.
- Deouell, L. Y., Heller, A. S., Malach, R., D’Esposito, M., & Knight, R. T. (2007). Cerebral responses to change in spatial location of unattended sounds. *Neuron*, 55(6), 985–996.
- Desimone, R., & Schein, S. J. (1987). Visual properties of neurons in area V4 of the macaque: Sensitivity to stimulus form. *Journal of Neurophysiology*, 57, 835–868.
- Desmurget, M., & Grafton, S. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, 4(11), 423–431.

- DeWitt, I., & Rauschecker, J. P. (2012). Phoneme and word recognition in the auditory ventral stream. *Proceedings of the National Academy of Sciences of the USA*, 109(8), E505–E514.
- Dhanjal, N. S., Handunnetthi, L., Patel, M. C., & Wise, R. J. (2008). Perceptual systems controlling speech production. *Journal of Neuroscience*, 28(40), 9969–9975.
- Diamond, I. T., Fisher, J. F., Neff, W. D., & Yela, M. (1956). Role of auditory cortex in discrimination requiring localization of sound in space. *Journal of Neurophysiology*, 19(6), 500–512.
- Ebeling, U., & von Cramon, D. (1992). Topography of the uncinate fascicle and adjacent temporal fiber tracts. *Acta Neurochirurgica*, 115(3–4), 143–148.
- Eliades, S. J., & Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology*, 89(4), 2194–2207.
- Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, 453(7198), 1102–1106.
- Engel, L. R., Frum, C., Puce, A., Walker, N. A., & Lewis, J. W. (2009). Different categories of living and non-living sound-sources activate distinct cortical networks. *NeuroImage*, 47(4), 1778–1791.
- Fecteau, S., Armony, J. L., Joannette, Y., & Belin, P. (2004). Is voice processing species-specific in the human brain? An fMRI study. *NeuroImage*, 23(3), 840–848.
- Frey, S., Campbell, J. S., Pike, G. B., & Petrides, M. (2008). Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *Journal of Neuroscience*, 28(45), 11435–11444.
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., & Anwander, A. (2006). The brain differentiates human and non-human grammars: Functional localization and structural connectivity. *Proceedings of the National Academy of Sciences of the USA*, 103(7), 2458–2463.
- Fu, K. G., Shah, A. S., Arnold, L., Garrahy, P. E., Smiley, J., Hackett, T. A., & Schroeder, C. E. (2003). Auditory cortical neurons respond to somatosensory stimulation. *Journal of Neuroscience*, 23, 7510–7515.
- Galaburda, A. M., Sanides, F., & Geschwind, N. (1978). Human brain: Cytoarchitectonic left-right asymmetries in the temporal speech region. *Archives of Neurology*, 35(12), 812–817.
- Gelfand, J. R., & Bookheimer, S. Y. (2003). Dissociating neural mechanisms of temporal sequencing and processing phonemes. *Neuron*, 38(5), 831–842.
- Geschwind, N. (1965). Disconnexion syndromes in animals and man. *Brain*, 88(2), 237–294, 585–644.
- Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *Journal of Neuroscience*, 25(20), 5004–5012.
- Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67–82). Hillsdale, NJ: Lawrence Erlbaum.
- Gifford, G. W., 3rd, & Cohen, Y. E. (2004). Effect of a central fixation light on auditory spatial responses in area LIP. *Journal of Neurophysiology*, 91(6), 2929–2933.
- Goldman-Rakic, P. S. (1996). The prefrontal landscape: Implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 351(1346), 1445–1453.
- Griffiths, T. D., Rees, A., Witton, C., Cross, P. M., Shakir, R. A., & Green, G. G. (1997). Spatial and temporal auditory processing deficits following right hemisphere infarction: A psychophysical study. *Brain*, 120(Pt 5), 785–794.
- Griffiths, T. D., Rees, G., Rees, A., Green, G. G., Witton, C., Rowe, D., et al. (1998). Right parietal cortex is involved in the perception of sound movement in humans. *Nature Neuroscience*, 1(1), 74–79.
- Grunewald, A., Linden, J. F., & Andersen, R. A. (1999). Responses to auditory stimuli in macaque lateral intraparietal area. I. Effects of training. *Journal of Neurophysiology*, 82(1), 330–342.
- Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27(3), 377–396; discussion 396–442.
- Guenther, F. H. (2006). Cortical interactions underlying the production of speech sounds. *Journal of Communication Disorders*, 39(5), 350–365.
- Hackett, T. A. (2011). Information flow in the auditory cortical network. *Hearing Research*, 271(1–2), 133–146.

- Hackett, T. A., Stepniewska, I., & Kaas, J. H. (1998). Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *Journal of Comparative Neurology*, 394(4), 475–495. 955
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, 9(7), 697–704. 958
- Hauser, M. D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press. 960
- Heffner, H., & Masterton, B. (1975). Contribution of auditory cortex to sound localization in the monkey (*Macaca mulatta*). *Journal of Neurophysiology*, 38(6), 1340–1358. 961
- Heffner, H. E., & Heffner, R. S. (1990). Effect of bilateral auditory cortex lesions on sound localization in Japanese macaques. *Journal of Neurophysiology*, 64(3), 915–931. 963
- Hershberger, W. (1976). Affference copy, the closed-loop analogue of von Holst's efference copy. *Cybernetics Forum*, 8, 97–102. 965
- Hickok, G., Okada, K., & Serences, J. T. (2009). Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *Journal of Neurophysiology*, 101(5), 2725–2732. 967
- Hikosaka, O., Nakahara, H., Rand, M. K., Sakai, K., Lu, X., Nakamura, K., et al. (1999). Parallel neural networks for learning sequential procedures. *Trends in Neurosciences*, 22(10), 464–471. 970
- Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *Journal of Cognitive Neuroscience*, 14(8), 1125–1138. 972
- Howard, M. A., Volkov, I. O., Mirsky, R., Garell, P. C., Noh, M. D., Granner, M., et al. (2000). Auditory cortex on the human posterior superior temporal gyrus. *Journal of Comparative Neurology*, 416, 79–92. 974
- Imig, T. J., Irons, W. A., & Samson, F. R. (1990). Single-unit selectivity to azimuthal direction and sound pressure level of noise bursts in cat high-frequency primary auditory cortex. *Journal of Neurophysiology*, 63(6), 1448–1466. 977
- Irvine, D. R. F. (1992). Physiology of auditory brainstem pathways. In R. R. Fay & A. A. Popper (Eds.), *The mammalian auditory pathway: Neurophysiology* (pp. 153–231). New York: Springer. 980
- Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J. L., S., Lin, F. H., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences of the USA*, 101(17), 6809–6814. 982
- Jenkins, W. M., & Merzenich, M. M. (1984). Role of cat primary auditory cortex for sound-localization behavior. *Journal of Neurophysiology*, 52(5), 819–847. 985
- Jones, E. G., Dell'Anna, M. E., Molinari, M., Rausell, E., & Hashikawa, T. (1995). Subdivisions of macaque monkey auditory cortex revealed by calcium-binding protein immunoreactivity. *Journal of Comparative Neurology*, 362(2), 153–170. 987
- Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences of the USA*, 97(22), 11793–11799. 990
- Kauramäki, J., Jääskeläinen, I. P., Hari, R., Möttönen, R., Rauschecker, J. P., & Sams, M. (2010). Transient adaptation of auditory cortex organization by lipreading and own speech production. *Journal of Neuroscience*, 30(4), 1314–1321. 992
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, 9(6), 718–727. 995
- Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2007). Functional imaging reveals visual modulation of specific fields in auditory cortex. *Journal of Neuroscience*, 27(8), 1824–1835. 997
- Keller, S. S., Roberts, N., & Hopkins, W. (2009). A comparative magnetic resonance imaging study of the anatomy, variability, and asymmetry of Broca's area in the human and chimpanzee brain. *Journal of Neuroscience*, 29(46), 14607–14616. 1000
- Kikuchi, Y., Horwitz, B., & Mishkin, M. (2010). Hierarchical auditory processing directed rostrally along the monkey's supratemporal plane. *Journal of Neuroscience*, 30(39), 13021–13030. 1002
- Korte, M., & Rauschecker, J. P. (1993). Auditory spatial tuning of cortical neurons is sharpened in cats with early blindness. *Journal of Neurophysiology*, 70(4), 1717–1721. 1004
- Krumbholz, K., Schönwiesner, M., Cramon, D. Y. v., Rübsamen, R., Shah, N. J., Zilles, K., & Fink, G. R. (2005). Representation of interaural temporal information from left and right auditory space in the human planum temporale and inferior parietal lobe. *Cerebral Cortex*, 15(3), 317–324. 1006

- 1010 Kusmirek, P., & Rauschecker, J. P. (2009). Functional specialization of medial auditory belt cor-
 1011 tex in the alert rhesus monkey. *Journal of Neurophysiology*, 102(3), 1606–1622.
- 1012 Kusmirek, P., Ortiz, M., & Rauschecker, J. P. (2012). Sound-identity processing in early areas of
 1013 the auditory ventral stream in the macaque. *Journal of Neurophysiology*, 107(4), 1123–1141.
- 1014 Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recog-
 1015 nition network while listening to newly acquired actions. *Journal of Neuroscience*, 27(2),
 1016 308–314.
- 1017 Lakatos, P., Chen, C. M., O'Connell, M. N., Mills, A., & Schroeder, C. E. (2007). Neuronal oscil-
 1018 lations and multisensory interaction in primary auditory cortex. *Neuron*, 53(2), 279–292.
- 1019 Leaver, A., Van Lare, J. E., Zielinski, B. A., Halpern, A., & Rauschecker, J. P. (2009). Brain activa-
 1020 tion during anticipation of sound sequences. *Journal of Neuroscience*, 29(8), 2477–2485.
- 1021 Leaver, A. M., & Rauschecker, J. P. (2010). Cortical representation of natural complex sounds:
 1022 Effects of acoustic features and auditory object category. *Journal of Neuroscience*, 30(22),
 1023 7604–7612.
- 1024 Lewis, J. W., & Van Essen, D. C. (2000). Corticocortical connections of visual, sensorimotor, and
 1025 multimodal processing areas in the parietal lobe of the macaque monkey. *Journal of Comparative*
 1026 *Neurology*, 428(1), 112–137.
- 1027 Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., & DeYoe, E. A. (2005). Distinct corti-
 1028 cal pathways for processing tool versus animal sounds. *Journal of Neuroscience*, 25(21),
 1029 5148–5158.
- 1030 Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of
 1031 the speech code. *Psychological Review*, 74, 431–461.
- 1032 Lomber, S. G., & Malhotra, S. (2008). Double dissociation of 'what' and 'where' processing in
 1033 auditory cortex. *Nature Neuroscience*, 11(5), 609–616.
- 1034 Lueschow, A., Miller, E. K., & Desimone, R. (1994). Inferior temporal mechanisms for invariant
 1035 object recognition. *Cerebral Cortex*, 4(5), 523–531.
- 1036 Maeder, P. P., Meuli, R. A., Adriani, M., Bellmann, A., Fornari, E., Thiran, J. P., et al. (2001).
 1037 Distinct pathways involved in sound recognition and localization: A human fMRI study.
 1038 *NeuroImage*, 14(4), 802–816.
- 1039 Malhotra, S., Hall, A. J., & Lomber, S. G. (2004). Cortical control of sound localization in the cat:
 1040 Unilateral cooling deactivation of 19 cerebral areas. *Journal of Neurophysiology*, 92(3),
 1041 1625–1643.
- 1042 Margoliash, D., & Fortune, E. S. (1992). Temporal and harmonic combination-sensitive neurons in
 1043 the zebra finch's HVc. *Journal of Neuroscience*, 12, 4309–4326.
- 1044 Mendelson, J. R., & Cynader, M. S. (1985). Sensitivity of cat primary auditory cortex (AI) neurons
 1045 to the direction and rate of frequency modulation. *Brain Research*, 327, 331–335.
- 1046 Meredith, M. A., & Clemo, H. R. (1989). Auditory cortical projection from the anterior ectosyl-
 1047 vian sulcus (field AES) to the superior colliculus in the cat: An anatomical and electrophysio-
 1048 logical study. *Journal of Comparative Neurology*, 289(4), 687–707.
- 1049 Merzenich, M. M., & Brugge, J. F. (1973). Representation of the cochlear partition on the superior
 1050 temporal plane of the macaque monkey. *Brain Research*, 50, 275–296.
- 1051 Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of
 1052 sound sequences in the auditory cortex of awake macaques. *Neuron*, 48(1), 139–148.
- 1053 Middlebrooks, J. C., Clock, A. E., Xu, L., & Green, D. M. (1994). A panoramic code for sound
 1054 location by cortical neurons. *Science*, 264(5160), 842–844.
- 1055 Morel, A., Garraghty, P. E., & Kaas, J. H. (1993). Tonotopic organization, architectonic fields, and
 1056 connections of auditory cortex in macaque monkeys. *Journal of Comparative Neurology*,
 1057 335(3), 437–459.
- 1058 Müller-Preuss, P., & Ploog, D. (1981). Inhibition of auditory cortical neurons during phonation.
 1059 *Brain Research*, 215(1–2), 61–76.
- 1060 Mullette-Gillman, O. A., Cohen, Y. E., & Groh, J. M. (2005). Eye-centered, head-centered, and
 1061 complex coding of visual and auditory targets in the intraparietal sulcus. *Journal of*
 1062 *Neurophysiology*, 94(4), 2331–2352.

- Narins, P. M., & Capranica, R. R. (1980). Neural adaptations for processing the two-note call of the Puerto Rican treefrog, *Eleutherodactylus coqui*. *Brain, Behavior and Evolution*, 17, 48–66.
- Numminen, J., Salmelin, R., & Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters*, 265(2), 119–122.
- Pandya, D. N., & Sanides, F. (1972). Architectonic parcellation of the temporal operculum in rhesus monkey and its projection pattern. *Zeitschrift für Anatomie und Entwicklungsgeschichte*, 139, 127–161.
- Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1999). Localization of cerebral activity during simple singing. *NeuroReport*, 10(18), 3979–3984.
- Petrides, M., & Pandya, D. N. (1984). Projections to the frontal cortex from the posterior parietal region in the rhesus monkey. *Journal of Comparative Neurology*, 228(1), 105–116.
- Petrides, M., & Pandya, D. N. (2009). Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *Public Library of Science Biology*, 7(8), e1000170.
- Pizzamiglio, L., Aprile, T., Spitoni, G., Pitzalis, S., Bates, E., D'Amico, S., & Di Russo, F. (2005). Separate neural systems for processing action- or non-action-related sounds. *NeuroImage*, 24(3), 852–861.
- Poremba, A., Saunders, R. C., Crane, A. M., Cook, M., Sokoloff, L., & Mishkin, M. (2003). Functional mapping of the primate auditory system. *Science*, 299(5606), 568–572.
- Poremba, A., Malloy, M., Saunders, R. C., Carson, R. E., Herscovitch, P., & Mishkin, M. (2004). Species-specific calls evoke asymmetric activity in the monkey's temporal poles. *Nature*, 427(6973), 448–451.
- Rajan, R., Aitkin, L. M., & Irvine, D. R. (1990a). Azimuthal sensitivity of neurons in primary auditory cortex of cats. II. Organization along frequency-band strips. *Journal of Neurophysiology*, 64(3), 888–902.
- Rajan, R., Aitkin, L. M., Irvine, D. R., & McKay, J. (1990b). Azimuthal sensitivity of neurons in primary auditory cortex of cats. I. Types of sensitivity and the effects of variations in stimulus parameters. *Journal of Neurophysiology*, 64(3), 872–887.
- Rauschecker, J. P. (1998a). Cortical processing of complex sounds. *Current Opinion in Neurobiology*, 8, 516–521.
- Rauschecker, J. P. (1998b). Parallel processing in the auditory cortex of primates. *Audiology and Neuro-Otology*, 3, 86–103.
- Rauschecker, J. P. (2005). Neural encoding and retrieval of sound sequences. *Annals of the New York Academy of Sciences*, 1060, 125–135.
- Rauschecker, J. P. (2007). Cortical processing of auditory space: Pathways and plasticity. In F. Mast & L. Jäncke (Eds.), *Spatial processing in navigation, imagery, and perception* (pp. 389–410). New York: Springer.
- Rauschecker, J. P. (2011). An expanded role for the dorsal auditory pathway in sensorimotor integration and control. *Hearing Research*, 271, 16–25.
- Rauschecker, J. P., & Korte, M. (1993). Auditory compensation for early blindness in cat cerebral cortex. *Journal of Neuroscience*, 13, 4538–4548.
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718–724.
- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proceedings of the National Academy of Sciences of the USA*, 97(22), 11800–11806.
- Rauschecker, J. P., & Tian, B. (2004). Processing of band-passed noise in the lateral auditory belt cortex of the rhesus monkey. *Journal of Neurophysiology*, 91(6), 2578–2589.
- Rauschecker, J. P., & Tian, B. (2005). Hierarchic processing of communication sounds in primates. In J. S. Kanwal & G. Ehret (Eds.), *Behavior and neurodynamics for auditory communication*. Cambridge, UK: Cambridge University Press.
- Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, 268(5207), 111–114.
- Rauschecker, J. P., Tian, B., Pons, T., & Mishkin, M. (1997). Serial and parallel processing in rhesus monkey auditory cortex. *Journal of Comparative Neurology*, 382, 89–103.

- 1117 Ravizza, R. J., & Masterton, B. (1972). Contribution of neocortex to sound localization in opossum
1118 (*Didelphis virginiana*). *Journal of Neurophysiology*, 35(3), 344–356.
- 1119 Reale, R. A., & Imig, T. J. (1980). Tonotopic organization in auditory cortex of the cat. *Journal of*
1120 *Comparative Neurology*, 192, 265–291.
- 1121 Recanzone, G. H. (2000). Spatial processing in the auditory cortex of the macaque monkey.
1122 *Proceedings of the National Academy of Sciences of the USA*, 97(22), 11829–11835.
- 1123 Recanzone, G. H., Guard, D. C., Phan, M. L., & Su, T. K. (2000). Correlation between the activity
1124 of single auditory cortical neurons and sound-localization behavior in the macaque monkey.
1125 *Journal of Neurophysiology*, 83(5), 2723–2739.
- 1126 Remedios, R., Logothetis, N. K., & Kayser, C. (2009). Monkey drumming reveals common net-
1127 works for perceiving vocal and nonvocal communication sounds. *Proceedings of the National*
1128 *Academy of Sciences of the USA*, 106(42), 18010–18015.
- 1129 Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic*
1130 *Bulletin & Review*, 12(6), 969–992.
- 1131 Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., & Behrens, T. E. (2008). The
1132 evolution of the arcuate fasciculus revealed with comparative DTI. *Nature Neuroscience*, 11(4),
1133 426–428.
- 1134 Rizzolatti, G., Ferrari, P. F., Rozzi, S., & Fogassi, L. (2006). The inferior parietal lobule: where
1135 action becomes perception. *Novartis Foundation Symposium*, 270, 129–140; discussion
1136 140–125, 164–129.
- 1137 Romanski, L. M., & Goldman-Rakic, P. S. (2002). An auditory domain in primate prefrontal cortex.
1138 *Nature Neuroscience*, 5(1), 15–16.
- 1139 Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker, J. P.
1140 (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal
1141 cortex. *Nature Neuroscience*, 2(12), 1131–1136.
- 1142 Sabes, P. N. (2000). The planning and control of reaching movements. *Current Opinion in*
1143 *Neurobiology*, 10(6), 740–746.
- 1144 Saur, D., Kreher, B. W., Schnell, S., Kummerer, D., Kellmeyer, P., Vry, M. S., et al. (2008). Ventral
1145 and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the*
1146 *USA*, 105(46), 18035–18040.
- 1147 Schubotz, R. I., Friederici, A. D., & von Cramon, D. Y. (2000). Time perception and motor timing:
1148 A common cortical and subcortical basis revealed by fMRI. *NeuroImage*, 11(1), 1–12.
- 1149 Scott, S. K., Blank, C. C., Rosen, S., & Wise, R. J. S. (2000). Identification of a pathway for intel-
1150 ligible speech in the left temporal lobe. *Brain*, 123, 2400–2406.
- 1151 Seltzer, B., & Pandya, D. N. (1994). Parietal, temporal, and occipital projections to cortex of the
1152 superior temporal sulcus in the rhesus monkey: A retrograde tracer study. *Journal of*
1153 *Comparative Neurology*, 343(3), 445–463.
- 1154 Smiley, J. F., Hackett, T. A., Ulbert, I., Karmas, G., Lakatos, P., Javitt, D. C., & Schroeder, C. E.
1155 (2007). Multisensory convergence in auditory cortex, I. Cortical connections of the caudal
1156 superior temporal plane in macaque monkeys. *Journal of Comparative Neurology*, 502(6),
1157 894–923.
- 1158 Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual
1159 inversion. *Journal of Comparative and Physiological Psychology*, 43(6), 482–489.
- 1160 Stricanne, B., Andersen, R. A., & Mazzoni, P. (1996). Eye-centered, head-centered, and intermediate
1161 coding of remembered sound locations in area LIP. *Journal of Neurophysiology*, 76(3),
1162 2071–2076.
- 1163 Suga, N., O'Neill, W. E., & Manabe, T. (1978). Cortical neurons sensitive to combinations of
1164 information-bearing elements of biosonar signals in the mustache bat. *Science*, 200, 778–781.
- 1165 Tata, M. S., & Ward, L. M. (2005a). Early phase of spatial mismatch negativity is localized to a
1166 posterior “where” auditory pathway. *Experimental Brain Research*, 167(3), 481–486.
- 1167 Tata, M. S., & Ward, L. M. (2005b). Spatial attention modulates activity in a posterior “where”
1168 auditory pathway. *Neuropsychologia*, 43(4), 509–516.
- 1169 Tian, B., & Rauschecker, J. P. (2004). Processing of frequency-modulated sounds in the lateral
1170 auditory belt cortex of the rhesus monkey. *Journal of Neurophysiology*, 92(5), 2993–3013.

- Tian, B., Reser, D., Durham, A., Kustov, A., & Rauschecker, J. P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science*, 292(5515), 290–293.
- Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *NeuroImage*, 39(3), 1429–1443.
- Von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip (Wechselwirkungen zwischen Zentralnervensystem und Peripherie). *Die Naturwissenschaften*, 37, 464–476.
- Wang, X. (2000). On cortical coding of vocal communication sounds in primates. *Proceedings of the National Academy of Sciences of the USA*, 97, 11843–11849.
- Warren, J. D., Zielinski, B. A., Green, G. G. R., Rauschecker, J. P., & Griffiths, T. D. (2002). Analysis of sound source motion by the human brain. *Neuron*, 34, 1–20.
- Warren, J. E., Wise, R. J., & Warren, J. D. (2005). Sounds do-able: Auditory-motor transformations and the posterior temporal plane. *Trends in Neurosciences*, 28(12), 636–643.
- Weeks, R. A., Aziz-Sultan, A., Bushara, K. O., Tian, B., Wessinger, C. M., Dang, N., et al. (1999). A PET study of human auditory spatial processing. *Neuroscience Letters*, 262(3), 155–158.
- Wernicke, C. (1874). *Der aphasische Symptomencomplex: Eine psychologische Studie auf anatomischer Basis*. Breslau: Cohn & Weigert.
- Wernicke, C. (1881). *Lehrbuch der Gehirnkrankheiten für Aerzte und Studierende*. Kassel, Berlin: Verlag Theodor Fischer.
- Wessinger, C. M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., & Rauschecker, J. P. (2001). Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, 13(1), 1–7.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7(7), 701–702.
- Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within 'Wernicke's area'. *Brain*, 124(Pt 1), 83–95.
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science*, 269(5232), 1880–1882.
- Wolpert, D. M., Doya, K., & Kawato, M. (2003). A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358(1431), 593–602.
- Yu, J. J., & Young, E. D. (2000). Linear and nonlinear pathways of spectral information transmission in the cochlear nucleus. *Proceedings of the National Academy of Sciences of the USA*, 97(22), 11780–11786.
- Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, 11(10), 946–953.
- Zatorre, R. J., Bouffard, M., Ahad, P., & Belin, P. (2002). Where is 'where' in the human auditory cortex? *Nature Neuroscience*, 5(9), 905–909.
- Zatorre, R. J., Bouffard, M., & Belin, P. (2004). Sensitivity to auditory object features in human temporal neocortex. *Journal of Neuroscience*, 24(14), 3637–3642.
- Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews Neuroscience*, 8(7), 547–558.
- Zimmer, U., & Macaluso, E. (2005). High binaural coherence determines successful sound localization and increased activity in posterior auditory areas. *Neuron*, 47(6), 893–905.



<http://www.springer.com/978-1-4614-2349-2>

Neural Correlates of Auditory Cognition

Cohen, Y.E.; Popper, A.N.; Fay, R.R. (Eds.)

2013, XIV, 330 p., Hardcover

ISBN: 978-1-4614-2349-2