

## The dual-pathway model of auditory signal processing

Wen-Jie WANG, Xi-Hong WU, Liang LI

*Department of Psychology, Speech and Hearing Research Center, Key Laboratory on Machine Perception, Ministry of Education, Peking University, Beijing 100871, China*

**Abstract:** Similar to the visual dual-pathway model, neurophysiological studies in non-human primates have suggested that the dual-pathway model is also applicable for explaining auditory cortical processing, including the ventral “what” pathway for object identification and the dorsal “where” pathway for spatial localization. This review summarizes evidence from human neuroimaging studies supporting the dual-pathway model for auditory cortical processing in humans.

**Keywords:** auditory perception; auditory localization; auditory pattern recognition; functional MRI

### 1 Introduction

**1.1 Auditory scene analysis** Perhaps the most intriguing question in auditory scene analysis is how listeners are able to identify, locate, and characterize individual sound sources<sup>[1]</sup>. Particularly in a cocktail party environment, parallel processing of both information of acoustic features and information of spatial locations of sound sources is critical for listeners. How the central auditory system processes “what” and “where” information is clearly an important issue in contemporary auditory neuroscience.

Before we review the recent literature associated with the existence of “what” and “where” auditory streams in human auditory cortex, we first briefly review the visual “what” and “where” dual-pathway model, because some interpretations of auditory anatomical and functional data are significantly influenced by this visual processing model.

**1.2 Visual “what” and “where” segregation pathways** It has been well known that different dimensions of visual information are processed by more than thirty visual cortical areas in the monkey cortex<sup>[2]</sup>, which compose two different functional streams<sup>[3,4]</sup>. The occipitotemporal pathway, or ven-

tral stream, is crucial for identifying objects, and the occipito-parietal pathway, or dorsal stream, is crucial for judging spatial locations and/or spatial relations among objects guiding movements in the space<sup>[3,5]</sup>. Specifically, neurons in the areas of the ventral stream selectively respond to shape, color, and texture<sup>[6]</sup> of the visual object, whereas neurons in the middle temporal area and other areas of the dorsal stream selectively respond to speed and direction of stimulus motions<sup>[7]</sup>.

Some striking similarities have been revealed between the human visual cortex and the monkey visual cortex by functional brain-imaging studies. Functional brain imaging techniques using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) can measure hemodynamic changes, blood flow, and blood oxygenation in cortical areas, as indirect measures of cortical neural activity<sup>[8,9]</sup>. The existence of separated processing streams has been examined via instructing human subjects to perform object-identity or spatial-localization tasks analogous to the tasks used in monkeys<sup>[10,11]</sup>. These studies show that regions in the ventral occipitotemporal cortex become more activated in the object identity tasks and regions in the dorsal occipito-parietal cortex become activated in the spatial-localization tasks, in agreement with results from studies using monkeys<sup>[12]</sup>.

Nevertheless, the neural computational strategies in the auditory system are quite different from those in the visual system. For example, acoustic spatial information is one-dimensionally conveyed by the motion of the tympanic

Corresponding author: Liang LI  
Tel: 86-10-62756804  
Fax: 86-10-62761081  
E-mail: liangli@pku.edu.cn  
Article ID: 1673-7067(2008)03-0173-010  
CLC number: Q426  
Document code: A  
Received date: 2007-12-26

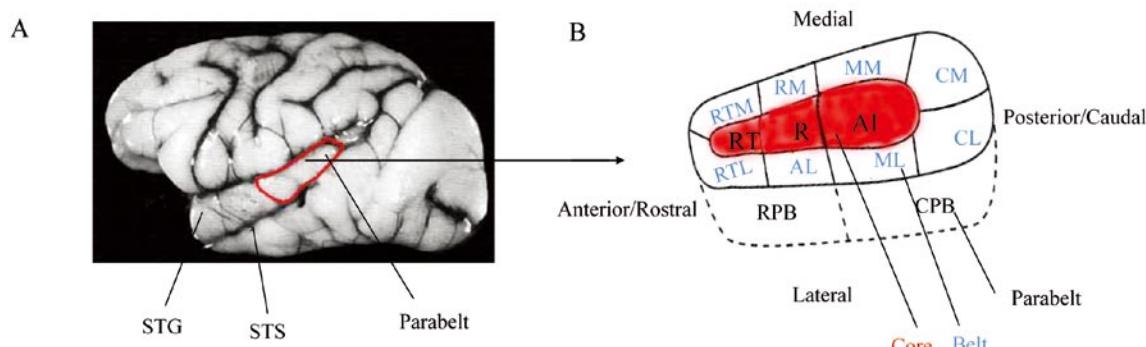
membrane, which leads to that auditory spatial information cannot be represented as a simple one-to-one mapping at the level of the peripheral receptors, but must be computed from binaural-difference cues, including the interaural level difference (ILD) and the interaural time difference (ITD).

Mammalian primary auditory cortex (PAC, or auditory area I; AI), which receives both monaural and binaural inputs from subcortical auditory nuclei as well as extensive interhemispheric projections, occupies a relatively later stage in the information-processing hierarchy than does the visual cortex, due to the greater anatomical complexity of the audi-

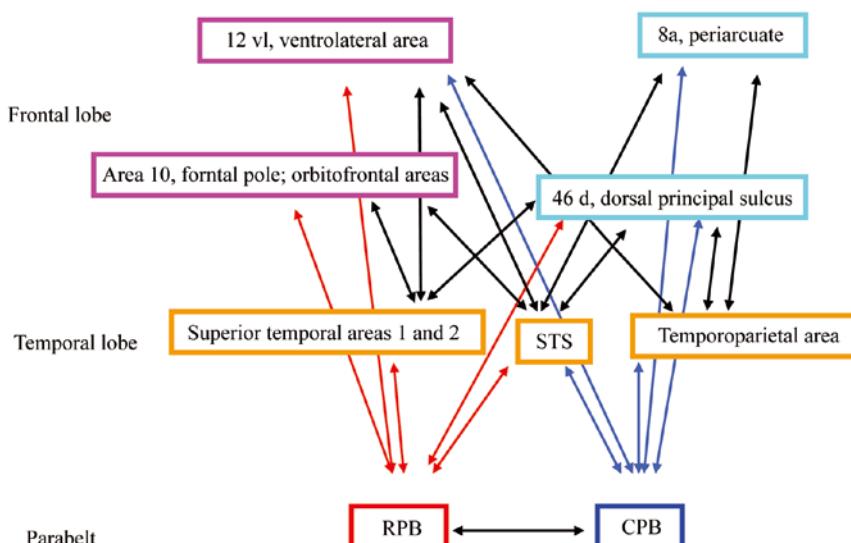
tory subcortical pathways. It has been suggested that certain synaptic properties and connectivity patterns of PAC are specific for extracting spectrotemporal feature information<sup>[13-15]</sup>. Hence, information processing of visual system and that of the auditory system may be quite different.

### 1.3 Auditory cortex studies in non-human primates

**1.3.1 The structural organization of auditory cortex** The auditory cortex in the macaque can be divided into a central core region surrounded by the belt and parabelt regions oriented along the long axis of the superior temporal lobe<sup>[16-20]</sup> (Fig. 1). In addition, the macaque auditory cortex has widely-



**Fig. 1** A, Lateral view of the macaque cerebral cortex. B, Enlarged view of the principal auditory areas in the macaque cerebral cortex. AI, auditory area I; AL, anterolateral area; CL, caudolateral area; CM, caudomedial area; CPB, caudal parabelt; ML, middlelateral area; R, rostral area; RM, rostromedial area; RPB, rostral parabelt; RT, rostrotemporal area; RTL, lateral rostrotemporal area; RTM, medial rostrotemporal area; STG, superior temporal gyrus; STS, superior temporal sulcus. Adapted from reference<sup>[18]</sup>.



**Fig. 2** Neural connections of the auditory cortex. Red and blue frames represent the rostral areas and the caudal areas of parabelt and frontal lobe, respectively. Correspondingly, red and blue arrows indicate the axonal connections projected from the rostral parabelt and caudal parabelt, respectively. The yellow frames represent the three main regions within the temporal lobe, and the black arrows represent the axonal connections projected from the temporal lobe.

distributed reciprocal extrinsic connections with rostral superior temporal gyrus (STG), insula, inferior parietal lobe (IPL), lateral prefrontal cortices, lateral amygdaloid nucleus, and subcortical structures including dorsal and medial divisions of the medial geniculate complex, putamen, inferior and superior colliculi (Fig. 2). The core, belt, parabelt, and their extrinsic connections can be considered as the various levels in the auditory processing hierarchy<sup>[18]</sup>.

The core is constituted with three fields: the most caudal AI, the most rostral field rostrotemporal area (RT), and the rostral area (R) between AI and RT<sup>[18]</sup> (for a review see<sup>[21]</sup>). Each of the core fields is characteristic of primary (or primary-like) sensory cortex. For examples, neurons in the core response to pure tones with the high frequency selectivity and short latencies, and all the core fields receive major axonal projections from the ventral nucleus of the medial geniculate body. Particularly, these core fields have well-developed layer 4 of granule cells which densely express both cytochrome oxidase, (the metabolic enzyme for deactivating the neurotransmitter acetylcholine) and parvalbumin (the calcium-binding protein). Unlike the primary visual cortex and the primary somatosensory cortex, the auditory core fields interconnect only with each other and with its neighbor, the belt, without long-distance projections. Thus the belt must be the obligatory second stage of auditory cortical processing, making the parabelt not to have direct access to core information. Each core field also connects reciprocally with several belt fields and the homotopic core field in the opposite hemisphere via the corpus callosum.

The belt, which has its dense connections with the core and the dorsal and medial divisions of the medial geniculate body, consists of eight fields with various representations of the cochlea. The densest projections to the caudal fields of the belt are from AI, and more rostral belt fields receive dense projections from R or RT. Unlike neurons in the core, neurons in belt fields respond better to narrow-band noises than pure tones, suggesting wider frequency-channel inputs to belt fields (for a review see<sup>[18]</sup>).

The rostral and caudal subdivisions of the parabelt, located between the rostral and caudal poles of the belt region, have connections with the belt region<sup>[18,21]</sup>. The parabelt also connects with a number of thalamic nuclei, including the dorsal and magnocellular divisions of the medial geniculate complex<sup>[22]</sup> and the medial pulvinar<sup>[18]</sup>. However, parabelt does

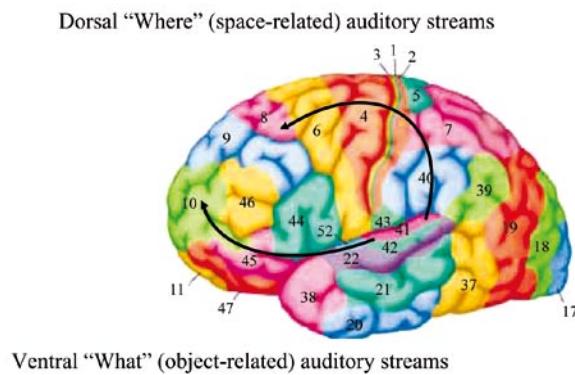
not have heavy connections with the core, and connections with the ventral medial geniculate are minimal, indicating that the belt may mediate auditory information flows from the core to the parabelt and even higher order cortical areas<sup>[21]</sup>.

Based on differences in connections, the parabelt has been divided into rostral and caudal divisions<sup>[18]</sup>. The rostral parabelt (RPB) receives projections mainly from rostral belt fields, whereas the caudal parabelt (CPB) receives projections mainly from caudal belt fields, and both divisions also receive projections from the rostromedial (RM) field of the belt.

The RPB projects to more rostral portions of the STG, whereas the CPB projects to the caudal end of the STG. The regions of the frontal lobe that receive projections from the parabelt include the frontal eyefield, the dorsolateral prefrontal cortex, the ventrolateral prefrontal cortex, and the orbital frontal cortex. Interestingly, based on the projection pattern of the parabelt to temporal and frontal cortex, it has been suggested that there are two both anatomically and functionally separate pathways for processing spatial and nonspatial auditory information<sup>[16,17]</sup>. Spatial and nonspatial regions of the prefrontal cortex are the targets of projections originating from caudal and rostral fields of non-primary (belt and parabelt) auditory cortex, respectively.

Thus, as summarized by Kaas and Hackett<sup>[18]</sup>, auditory cortical signal processes can be divided into four stages, which occur in the core, belt, parabelt, and targets of parabelt projections, respectively. The formation of the “what” and the formation of “where” pathways (see below) have their roots in these four-stage processes.

**1.3.2 “What” and “where” streams** In electrophysiological studies, when comparing the spatial selectivity of single neurons across the caudal and rostral belts and the parabelt region of rhesus monkeys, neurons located in caudolateral areas are the most sharply tuned and those in rostralateral areas are most broadly tuned. However, when comparing the selectivity to “what” signals, such as monkey calls, neurons located in rostralateral areas are more selective than those in caudolateral areas<sup>[23-26]</sup>. These relative differences in stimulus selectivity have lead to the formation of the concept of dual ventral-“what” (object-related) and dorsal-“where” (space-related) auditory-processing streams, with separated projections arising in the caudal and rostral superior temporal plane (STP) to the spatial (i.e., areas 8a, caudal 46d) and



**Fig. 3** Prefrontal connections of auditory cortex in macaque monkeys.  
Adapted from reference<sup>[19]</sup>.

nonspatial (i.e., areas 10,12vl, rostral 46d) areas in prefrontal cortex<sup>[16, 17, 19, 27, 28]</sup> (Fig. 3).

## 2 “What” and “where” pathways in human auditory cortex

Here we start to review the recently published literatures related with investigations of the existence of “what” and “where” auditory processing streams in human cerebral cortex, including non-invasive studies using fMRI, electrical event-related potentials (ERPs), magnetoencephalography (MEG), and cortical lesion methods. MEG measures the weak magnetic fields produced by electric currents flowing in neurons, and is a useful non-invasive technique for investigating human cerebral cortex activity with the time resolution better than 1 ms and the spatial resolution of 2-3 mm<sup>[29]</sup>.

**2.1 Human auditory cortex anatomy** Human auditory cortex shares the similar organizational scheme with the macaque and other primates, consisting of a central primary area, the core ( Brodmann areas 41 and 42; Heschl’s gyrus, HG), which is tonotopically organized and surrounded by multiple non-primary fields (Brodmann area 22, planum temporale: PT)<sup>[30-34]</sup> (for a review see<sup>[18]</sup>; Fig. 4).

### 2.2 Non-invasive studies

**2.2.1 Functional imaging studies** Using identical noise bursts in both the pitch and spatial discrimination tasks with differing instructions to human subjects, Alain *et al.* assessed participants’ regional cerebral activity associating with the pitch task and spatial discrimination task, respectively<sup>[35]</sup>. By pressing one of three buttons, participants indicated whether the second noise burst (S2) was lower, identical, or higher in pitch than the first noise burst (S1), regardless of its location;



**Fig. 4** Schematic drawings of the horizontal view of principal auditory areas in the human left STG. Adapted from reference<sup>[34]</sup>.

or S2 was at a leftward, an identical, or a rightward position relative to S1, regardless of changes in pitch. They directly compared the changes in hemodynamic responses obtained during the pitch task with changes in the responses obtained during the location task. Primary auditory cortices, extending anteriorly to auditory association cortices on STP and the right inferior frontal gyrus had greater activation associated with the pitch judgment. Conversely, selectively processing sound locations was associated with the enhanced blood-oxygen-level dependent (BOLD) signal in bilateral posterior temporal areas, inferior and superior parietal cortices compared with the pitch judgment. This study provides the direct evidence that the neural systems involved in identifying and localizing auditory objects are functionally and neuroanatomically segregated based on task demands even when stimuli are identical across tasks.

Nevertheless, studies in this line of dissociation of “where” and “what” processing have adopted working memory paradigms that require sound information to be perceived, temporarily stored, reorganized, and then compared with a test stimulus. Thus the spatial- and nonspatial-processing dissociation in neural pathways may result from numerous cognitive processes. The dissociation can take place at the time intervals of slow evoked responses of P3 and positive slow wave (PSW): the occipito-temporal generator of the P3 is activated more strongly during the performance of the location task, while the activity of the temporal generator of the PSW is enhanced during working memory processing of sound pitch<sup>[36]</sup>. In addition, the “where”-ver-

sus-“what” dissociation pattern in working memory tasks varies across task stages, such as the sample, delay, and test periods as used in study of Rämä *et al.*<sup>[37]</sup>, and the early and later stages used in study of Arnott *et al.*<sup>[38]</sup>. Furthermore, the dissociation may not be specific to certain types of acoustic stimuli, but reflect a general spatial versus nonspatial distinction. For example, similar dissociation was observed by using band-pass noise<sup>[35,38]</sup> and using speech samples<sup>[37]</sup>.

Maeder *et al.*<sup>[39]</sup> also found that comparison of activation by recognition versus localization with activation by localization versus recognition show the segregated pathways involved in sound recognition and those in sound localization. In the recognition task, participants were asked to press a button when they recognize the animal cries under the background of everyday environment, and in the localization task they were asked to discriminate whether two 500-ms-long white noise tracts under the white noise background were presented at the same or at two different locations within the same hemifield. Their results show that bilateral middle temporal gyrus (MTG), precuneus, and the left posterior part of inferior frontal gyrus were more activated by meaningful, environmental sounds recognition than white noise tracts sound localization; while lower part of IPL, posterior parts of middle and inferior frontal gyri were more activated bilaterally by sound localization than sound recognition. Passive listening to stimuli also yielded distinct activation patterns: MTG and posterior prefrontal cortex on both sides were more activated by listening to meaningful than to spatialized sounds, while the inferior part of IPL and the posterior part of STG were more activated on both sides by listening to spacialized than to meaningful sounds<sup>[39]</sup>, which indicated that the differential activation of the two pathways resulting from an organizational principle of human auditory cortex rather than the possible attention or motor biases involved in the active task.

A large region involving STP posterior to HG in human auditory association cortex belongs to PT, which has been suggested to be crucial in distinguishing different types of sound patterns, including auditory spatial information and sound object spectrotemporal features<sup>[40,41]</sup>. PT may play a role in gating auditory information between the “what” and “where” streams. Using the fMRI method, Warren *et al.* demonstrated that some subregions within human PT process particular sound attributes<sup>[34]</sup>. In this study, contrasts be-

tween conditions with changing pitch and fixed pitch produced bilateral activation involving lateral HG, anterior planum PT, and planum polare (PP) anterior to HG, extending into STG, while contrasts between all conditions with changing spatial location and fixed location produced bilateral activation involving posterior PT, which is bilaterally posterior to those for pitch change.

However, using the fMRI method, Barret *et al.* found that the pitch-related activity in PT also overlapped with the response to the spatially shifting sound source, suggesting that anterior PT is not specialized for spatial or non-spatial processing<sup>[42]</sup>. They manipulated both the pitch feature (no-pitch noise, fixed-pitch noise, or varied-pitch noise) and the spatial feature of noise stimuli. Compared to no-pitch noise bursts, pitch noises produced widespread bilateral activation within and anterior to HG, PP, and the anterior part of PT. Compared to diffuse sound sources, compact sound sources evoked activation in bilateral PT, which was different from activation associated with pitch. They also found that the different spatial information were not equally associated with the “where” stream: the posterior auditory cortex prefer responding to location shifts of compact sounds rather than changes of spatial width.

**2.2.2 Interaction of “what” and “where” processing** Some visual recognition theories suggest that through bottom-up and top-down interactions, the faster “where” visual pathway may provide earlier rough object identity assessment for the slower and more detailed object identification in “what” pathway<sup>[43]</sup>. An early top-down projection from the orbitofrontal cortex to the visual cortex may facilitate the object recognition. Electrical physiological studies have also demonstrated the latency advantage of the dorsal over the ventral visual pathway in Macaque<sup>[44]</sup>.

Human psychophysical studies have indicated that the interaction of “where” pathway and “what” pathway must be existent in the auditory system. When two highly correlated sounds are presented by two spatially separated loudspeakers, if the onset delay between the two sounds is sufficiently short, the listener perceives only one fused sound image as coming from the leading loudspeaker without perceiving a second distinct sound image as coming from the lagging loudspeaker. This perceptual fusion is certainly related to auditory spatial processing and can be used for inducing perceived spatial separation of uncorrelated sounds.

For example, when two uncorrelated sounds A and B are presented by a loudspeaker to the listener's left and a loudspeaker to the right, and the left loudspeaker leads the right loudspeaker for sound A by 3 ms and the right loudspeaker leads the left loudspeaker for sound B by 3 ms, the listener perceives sound A as only coming from a location near the left loudspeaker and sound B as only coming from a location near the right loudspeaker. Thus the sound-A image and the sound-B image are spatially separated, even though both sounds are physically delivered by each of the two loudspeakers. Interestingly, when sound A is used as the masker and sound B is used as the target signal, although perceived spatial separation between A (masker) and B (signal) does not substantially changes acoustics at the listener's ears, it significantly improves (unmasks) recognition of the signal, indicating that the fusion-related spatial processing affects perception of non-spatial features of the target signal<sup>[45,46]</sup>. Moreover, when the target signal is speech, unmasking effect caused by perceived spatial separation is much larger when the masker is speech than when the masker is steady-state speech-spectrum noise. The results indicate that the unmasking effect of perceived spatial separation is modulated by perceptual, cognitive, and/or even linguistic processing of non-spatial features of speech. We propose that the interaction between perceived spatial separation and speech recognition in the presence of maskers can be used as a new behavioral model for brain-imaging studies of the interaction of auditory "where" and "what" pathways.

MEG studies<sup>[47]</sup> in humans have shown that there are two auditory cortical sources that contribute to the N1 response: an earlier N1 that originates in the posterior auditory cortex and is less sensitive to sound frequencies, and a later N1 that originates in the anterior auditory cortex and is narrowly tuned to sound frequencies. In addition, with decreasing sound novelty by repeated sound presentations, the amplitude of the posterior N1 source is rapidly suppressed, whereas that of the anterior N1 source is less affected. It is suggested that the posterior auditory cortex, which is associated with "where" processing, is important for fast analysis of novel sound location and attentional orientation. On the other hand, the anterior auditory cortex is more involved in subsequent attentional analysis of the fine-object features<sup>[47]</sup>.

Similar to the visual system, the interaction between the

auditory "where" and "what" pathways reflects the brain processing strategy of using the interaction between bottom-up and the top-down signaling in complex environments. In the visual system, fast signal-driven (bottom-up) pathways activate the left orbitofrontal cortex to process low spatial-frequency information for coarse "initial guesses" of object identifying, and the orbitofrontal cortex in turn top-down modulates the recognition-related areas in the temporal cortex, allowing both attention to be shifted to the target object and precise processing of detailed non-spatial features of the target object<sup>[43]</sup>. In the auditory system, spatial information of a novel acoustic event is rapidly and coarsely processed in the posterior auditory to shift and maintain attention to the relevant auditory object, allowing the "what" pathways to process detailed non-spatial information.

However, the neural circuits of the feedback projections and its temporal process have not been clearly demonstrated in human cerebral cortex. Our knowledge to this interaction needs further extension. MEG measurements have shown that the posterior auditory "where" stream is activated significantly earlier (about 30 ms) than the anterior "what" stream<sup>[48]</sup>. This study also demonstrated that the human non-primary auditory cortex processes speech-sound identity and location in parallel with anterior "what" (in anterolateral HG, anterior STG, and posterior PP) and posterior "where" (in PT and posterior STG) pathways as early as 70 to 150 ms from the stimulus onset. These results are supported by a recent ERP study<sup>[49]</sup>.

One fMRI-constrained ERP source analysis also indicates that location changes are processed faster than pattern changes by about 100 ms<sup>[50]</sup>. The location change condition (using head-related transfer function) elicits activity in the posterior temporal area PT and posterior superior temporal area about 300 ms after the stimulus onset, while the pattern change (two different animal vocalizations) elicits a broader and more pronounced dipole response in the more anterior parts of the temporal cortex, i.e., anterior STG and superior temporal sulcus (STS), PP, lateral HG and anterior PT, about 400-500 ms after the stimulus onset.

**2.2.3 Cross-modal studies** The segregation between object recognition and spatial processing seems to be a general principle of the cortical function organization, since the what/ventral-where/dorsal model is also valid in the cross-modal audiovisual stimulation. Sestieri *et al.* used both visual and

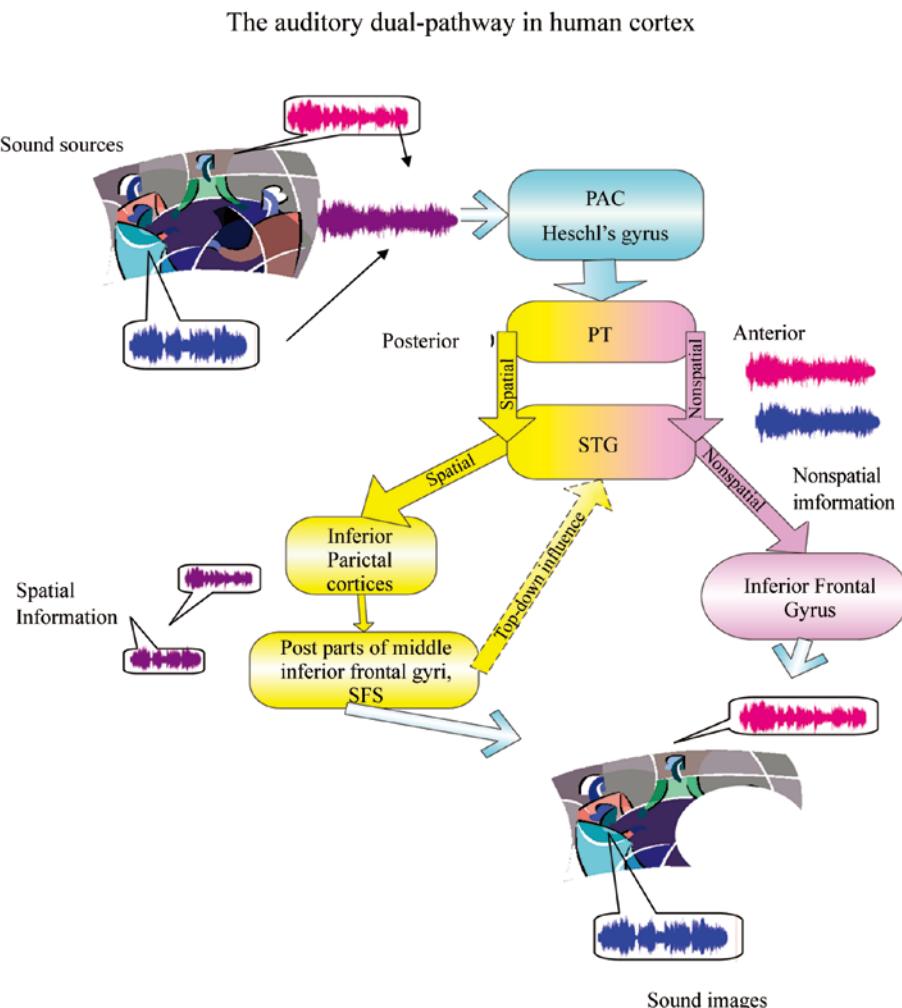
auditory stimuli for performing either a semantic content recognition or a spatial position localization matching task<sup>[51]</sup>. The fMRI results show that the cross-modal localization task elicited more activity than the recognition task in bilateral precuneus, right IPL, right intraparietal sulcus, and the right superior occipital cortex. Regions responding to the cross-modal recognition task more than the localization task were found in the bilateral inferior occipital gyrus and the left lateral temporal cortex including the anterior part of STS and STG.

**2.3 Lesion studies** Clinical disorders of central auditory function have also shown that the dissociation between sound localization and sound recognition associated with distinct cerebral lesion regions. Patients with severely deficient in recognition of environmental sounds but normal in

auditory localization had the lesion in some regions, including the left superior, middle and inferior temporal gyri, lateral auditory areas, and the anterior parts of the temporal lobe. However, severely deficient in auditory motion perception and partially deficient in auditory localization, but normal in recognition of environmental sounds, was accompanied with damage to a dorsal temporo-parieto-frontal region<sup>[52,53]</sup>.

### 3 Summary

The dual-pathway auditory-processing model is summarized in Fig. 5. Based on the existing literatures mentioned in this review article, it is clear that the segregation as well as the interaction of the spatial-feature processing and non-spatial-feature processing reflects one of most important functional organizations of human auditory cortex. Here we pro-



**Fig. 5** The dual-pathway model for explaining auditory processing in human cortex. The wave color of red and blue indicate the different pitch information. The yellow pathway indicates the spatial information pathway. The pink pathway indicates the nonspatial information pathway.

pose that this functional organization is critical for facilitating auditory scene analyses under noisy, even reverberant environments.

Under such adverse environments, the auditory system has to be able to differentiate sound waves from various sources, selectively analyze the target signals, and suppress irrelevant stimuli. Due to the limited capacity for deeper informational processing, the auditory system needs to set up a target signal among complex signal sources before it carries out processing of detailed acoustic features. Sound localization appears to be the first processing stage. According to the model illustrated in Fig. 5, spatial information of acoustic objects is proposed to be projected from the primary auditory cortex into PT, then transferred to IPL, post parts of middle inferior frontal gyri, and superior frontal regions. It is well known that sound localization largely depends on auditory processing of interaural time differences. Binaural timing information can be translated into region-specific activity in the auditory cortex. More specifically, the interaural temporal processing of lateralized sounds elicits responses in the contralateral PT, and when the interaural time cue changes in association with a movement, the responses become stronger and extend further into adjacent regions of the IPL<sup>[54]</sup>. The involvement of IPL in processing of spatial information implies that attention starts to play a role in the task, because it has been confirmed that the multiple-modal IPL is important in processing visual spatial information and orienting visual attention<sup>[55,56]</sup>. Indeed, when the auditory task requires listeners to selectively compare or evaluate the location of sound source, IPL activity is specifically enhanced<sup>[54,57,58]</sup>, confirming that the parietal cortex also plays an important role in modulating auditory spatial attention. To further enhance the listener's attention to the target signal, the prefrontal cortex becomes activated through the fast "where" pathway originating in the auditory cortex, and then projects to STG to facilitate the sound pattern analyses of non-spatial "what" information. Under the prefrontal top-down modulation, the anterior portion of auditory cortex (including anterior STG and MTG, and STS) has been further activated in analyses of certain types of spectrotemporal pattern, from simple spectral variance to voices and speech<sup>[59-61]</sup>, allowing the brain to further encode target-signal details (for a meta-analysis see<sup>[62]</sup>). Finally, co-activation of the anterior temporal cortex and the inferior frontal cortex are critical for

conscious perception of the target at both global and local levels. Because selective attention plays a role in this interaction between "where" and "what" processing, irrelevant signals (masking signals) in the environment are suppressed at the same time. As mentioned, recent studies have confirmed that perceived spatial separation between target speech and masker can allow listener to selectively attend to target speech and ignore masking stimuli, leading to a release of target speech from masking, particularly from speech masking<sup>[45,46]</sup>.

It should be noted that the number of auditory processing streams may not be limited only to two. There could be several 'streamlets' involved in various aspects of auditory perception, such as a dorsal "how" pathway proposed to computing the spatiotemporal modulation of sounds<sup>[63]</sup>. Clearly, future investigations are needed to enrich our knowledge of the functional organizations of the sensory/perception systems of the brain.

**Acknowledgements:** This work was supported by the National Natural Science Foundations of China (No. 30711120563, No. 30670704, and No. 60535030).

## References:

- [1] Bregman AS. Auditory Scene Analysis. Cambridge, MA: MIT Press, 1990.
- [2] Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1991, 1: 1-47.
- [3] Ungerleider LG, Mishkin M. Two cortical visual systems. In: Ingle D, Goodale M, Mansfield R Ed. *Analysis of Visual Behavior*. Cambridge, MA: MIT Press, 1982, 49-86.
- [4] Ungerleider LG. Functional brain imaging studies of cortical mechanisms for memory. *Science* 1995, 270: 769-775.
- [5] Goodale MA, Milner AD. Separate visual pathways for perception and action. *Trends Neurosci* 1992, 15: 20-25.
- [6] Desimone R, Ungerleider LG. Neural mechanisms of visual processing in monkeys. In: Boller F, Grafman J Ed. *Handbook of Neuropsychology*. Amsterdam: Elsevier, 1989, 267-299.
- [7] Goldberg ME, Colby C. The neurophysiology of spatial vision. In: Boller F, Grafman J Ed. *Handbook of Neuropsychology*. Amsterdam: Elsevier, 1989, 301-315.
- [8] Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, et al. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc Natl Acad Sci* 1992, 89: 5675-5679.
- [9] Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, et al. Intrinsic signal changes accompanying sensory stimulation:

- functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci 1992, 89: 5951-5955.
- [10] Haxby JV, Horwitz B, Ungerleider LG, Maisog JM, Pietrini P, Grady CL. The functional organization of human extrastriate cortex: a PET-rCBF study of selective attention to faces and locations. J Neurosci 1994, 14: 6336-6353.
- [11] Ungerleider LG, Haxby JV. "what" and "where" in the human brain. Curr Opin Neurobiol 1994, 4: 157-165.
- [12] Kastner S, Ungerleider LG. Mechanisms of visual attention in the human cortex. Annu Rev Neurosci 2000, 23: 315-341.
- [13] Atzori M, Lei S, Evans DI, Kanold PO, Phillips-Tansey E, McIntyre O, et al. Differential synaptic processing separates stationary from transient inputs to the auditory cortex. Nat Neurosci 2001, 4: 1230-1237.
- [14] Miller LM, Escabi MA, Read HL, Schreiner CE. Functional convergence of response properties in the auditory thalamocortical system. Neuron 2001, 32: 151-160.
- [15] Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci 2002, 6: 37-46.
- [16] Romanski LM, Tian B, Fritz JB, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. Nat Neurosci 1999, 2: 1131-1136.
- [17] Romanski LM, Bates JF, Goldman-Rakic PS. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. J Comp Neurol 1999, 403: 141-157.
- [18] Kaas JH, Hackett TA. Subdivisions of auditory cortex and processing streams in primates. Proc Natl Acad Sci 2000, 97: 11793-11799.
- [19] Poremba A, Saunders RC, Crane AM, Cook M, Sokoloff L. Functional mapping of the primate auditory system. Science 2003, 299: 568-572.
- [20] Jones EG, Dellanna ME, Molinari M, Rausell E, Hashikawa T. Subdivisions of macaque monkey auditory cortex revealed by calcium-binding protein immunoreactivity. J Comp Neurol 1995, 362: 153-170.
- [21] Hackett TA, Stepniewska I, Kaas JH. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 1998, 394: 475-495.
- [22] Hackett TA, Stepniewska I, Kaas JH. Thalamocortical connections of the parabelt auditory cortex in macaque monkeys. J Comp Neurol 1998, 400: 271-286.
- [23] Rauschecker JP. Parallel processing in the auditory cortex of primates. Audiol Neurotol 1998, 3: 86-103.
- [24] Rauschecker JP, Tian B. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci 2000, 97: 11800-11806.
- [25] Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. Functional specialization in rhesus monkey auditory cortex. Science 2001, 292: 290-293.
- [26] Woods TM, Lopez SE, Long JH, Rahman JE, Recanzone GH. Effects of stimulus azimuth and intensity on the single-neuron activity in the auditory cortex of the alert macaque monkey. J Neurophysiol 2006, 96: 3323-3337.
- [27] Kaas JH, Hackett TA. "what" and "where" processing in auditory cortex. Nat Neurosci 1999, 2: 1045-1047.
- [28] Romanski LM, Tian B, Fritz JB, Mishkin M, Goldman-Rakic PS, Rauschecker JP. Reply to "what", "where" and "how" in auditory cortex. Nat Neurosci 2000, 3: 966.
- [29] Hamalainen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. Reviews of Modern Physics 1993, 65: 413-497.
- [30] Hackett TA, Preuss TM, Kaas JH. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. J Comp Neurol 2001, 441: 197-222.
- [31] Fullerton BC, Pandya DN. Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. J Comp Neurol 2007, 504: 470-498.
- [32] Wallace MN, Johnston PW, Palmer AR. Histochemical identification of cortical areas in the auditory region of the human brain. Exp Brain Res 2002, 143: 499-508.
- [33] Ratnanather JT, Barta PE, Honeycutt NA, Lee N, Morris HM, Dziorny AC, et al. Dynamic programming generation of boundaries of local coordinatized submanifolds in the neocortex: application to the planum temporale. Neuroimage 2003, 20: 359-377.
- [34] Warren JD, Griffiths TD. Distinct mechanisms for processing spatial sequences and pitch sequences in the human auditory brain. J Neurosci 2003, 23: 5799-5804.
- [35] Alain C, Arnott SR, Hevenor S, Graham S, Grady CL. "what" and "where" in the human auditory system. Proc Natl Acad Sci 2001, 98: 12301-12306.
- [36] Anurova I, Artchakov D, Korvenoja A, Ilmoniemi RJ, Aronen HJ, Carlson Sv. Cortical generators of slow evoked responses elicited by spatial and nonspatial auditory working memory tasks. Clin Neurophysiol 2005, 116: 1644-1654.
- [37] Rämä P, Poremba A, Sala JB, Yee L, Malloy M, Mishkin M, et al. Dissociable functional cortical topographies for working memory maintenance of voice identity and location. Cereb Cortex 2004, 14: 768-780.
- [38] Arnott SR, Grady CL, Hevenor SJ, Graham S, Alain C. The functional organization of auditory working memory as revealed by fMRI. J Cogn Neurosci 2005, 17: 819-831.
- [39] Maeder PP, Meuli RA, Adriani M, Bellmann A, Fornari E, Thiran JP, et al. Distinct pathways involved in sound recognition and localization: A human fMRI study. Neuroimage 2001, 14: 802-816.
- [40] Warren JD, Zielinski BA, Green GG, Rauschecker JP, Griffiths TD. Perception of sound-source motion by the human brain. Neuron

- 2002, 34: 139-148.
- [41] Griffiths TD, Warren JD. The planum temporale as a computational hub. *Trends Neurosci* 2002, 25: 348-353.
- [42] Barrett DJK, Hall DA. Response preferences for "what" and "where" in human non-primary auditory cortex. *Neuroimage* 2006, 32: 968-977.
- [43] Bar M, Kassam KS, Ghuman AS, Boshyan J, Schmid AM, Dale AM, et al. Top-down facilitation of visual recognition. *Proc Natl Acad Sci* 2006, 103: 449-454.
- [44] Chen CM, Lakatos P, Shah AS, Mehta AD, Givre SJ, Javitt DC, et al. Functional anatomy and interaction of fast and slow visual pathways in macaque monkeys. *Cereb Cortex* 2007, 17: 1561-1569.
- [45] Freyman RL, Helfer KS, McCall DD, Clifton RK. The role of perceived spatial separation in the unmasking of speech. *J Acoust Soc Am* 1999, 106: 3578-3588.
- [46] Wu Y, Li W, Chen J, Wang C, Qu H, Wu X, et al. Informational masking of Chinese speech under perceived spatial separation. *Chinese J. Acoustics (in Chinese)* 2005, 30: 462-467.
- [47] Jääskeläinen IP, Ahveninen J, Bonmassar G, Dale AM, Ilmoniemi RJ, Levänen S, et al. Human posterior auditory cortex gates novel sounds to consciousness. *Proc Natl Acad Sci* 2004, 101: 6809-6814.
- [48] Ahveninen J, Jääskeläinen IP, Raij T, Bonmassar G, Devore S, Hämaläinen M, et al. Task-modulated "what" and "where" pathways in human auditory cortex. *Proc Natl Acad Sci* 2006, 103: 14608-14613.
- [49] De Santis L, Clarke S, Murray MM. Automatic and intrinsic auditory "what" and "where" processing in humans revealed by electrical neuroimaging. *Cereb Cortex* 2007, 17: 9-17.
- [50] Altmann CF, Bledowski C, Wibral M, Kaiser J. Processing of location and pattern changes of natural sounds in the human auditory cortex. *Neuroimage* 2007, 35: 1192-1200.
- [51] Sestieri C, Di Matteo R, Ferretti A, Del Gratta C, Caulo M, Tartaro A, et al. "What" versus "where" in the audiovisual domain: An fMRI study. *Neuroimage* 2006, 33: 672-680.
- [52] Clarke S, Bellmann A, Meuli RA, Assal G, Steck AJ. Auditory agnosia and auditory spatial deficits following left hemispheric lesions: evidence for distinct processing pathways. *Neuropsychologia* 2000, 38: 797-807.
- [53] Clarke S, Bellmann Thiran A, Maeder P, Adriani M, Vernet O, Regli L, et al. What and where in human audition: selective deficits following focal hemispheric lesions. *Exp Brain Res* 2002, 147: 8-15.
- [54] Krumbholz K, Schönwiesner M, von Cramon DY, Rübsamen R, Shah NJ, Zilles K, et al. Representation of interaural temporal information from left and right auditory space in the human planum temporale and inferior parietal lobe. *Cereb Cortex* 2005, 15: 317-324.
- [55] Rushworth MF, Krams M, Passingham RE. The attentional role of the left parietal cortex: The distinct lateralization and localization of motor attention in the human brain. *J Cogn Neurosci* 2001, 13: 698-710.
- [56] Federspiel A, Volpe U, Horn H, Dierks T, Franck A, Vannini P, et al. Motion standstill leads to activation of inferior parietal lobe. *Hum Brain Mapp* 2006, 27: 340-349.
- [57] Griffiths TD, Rees G, Rees A, Green GG, Witton C, Rowe D, et al. Right parietal cortex is involved in the perception of sound movement in humans. *Nat Neurosci* 1998, 1: 74-79.
- [58] Bushara KO, Weeks RA, Ishii K, Catalan MJ, Tian B, Rauschecker JP, et al. Modality-specific frontal and parietal areas for auditory and visual spatial localization in humans. *Nat Neurosci* 1999, 2: 759-766.
- [59] Warren JD, Jennings AR, Griffiths TD. Analysis of the spectral envelope of sounds by the human brain. *Neuroimage* 2005, 24: 1052-1057.
- [60] Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B. Voice-selective areas in human auditory cortex. *Nature* 2000, 403: 309-312.
- [61] Specht K, Reul J. Functional segregation of the temporal lobes into highly differentiated subsystems for auditory perception: an auditory rapid event-related fMRI-task. *Neuroimage* 2003, 20: 1944-1954.
- [62] Arnott SR, Binns MA, Grady CL, Alain C. Assessing the auditory dual-pathway model in humans. *Neuroimage* 2004, 22: 401-408.
- [63] Belin P, Zatorre RJ. "What", "where" and "how" in auditory cortex. *Nat Neurosci* 2000, 3: 965-966.

## 听觉信号加工的双通路模型

王雯洁，吴玺宏，李量

北京大学心理学系，言语和听觉研究中心，视觉听觉国家教育部重点实验室，北京 100871

**摘要：**神经电生理学研究发现，处理视觉信息的人大脑皮层的双通路模型概念也适用于非人灵长类动物大脑皮层听觉信息的加工，即存在一条识别物体特征的腹侧通路和一条加工空间信息的背侧通路。本文系统地总结了近年来有关人类听觉加工的脑成像研究，并根据这些实验结果对人听觉双通路的加工理论模型进行了讨论。

**关键词：**听觉感知；听觉定位；听觉物体识别；核磁共振功能成像