

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^[1]. 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.

It 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 influence

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 differ-

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^[18].

The core is constituted with three fields: the most caudal AI, the most rostral field rostromedial temporal area (RT), and the rostral area (R) between AI and RT^[18] (for a review see^[21]). Each of the core fields is characteristic of primary (or primary-like) sensory cortex. For examples, neurons in the core respond 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-devel-

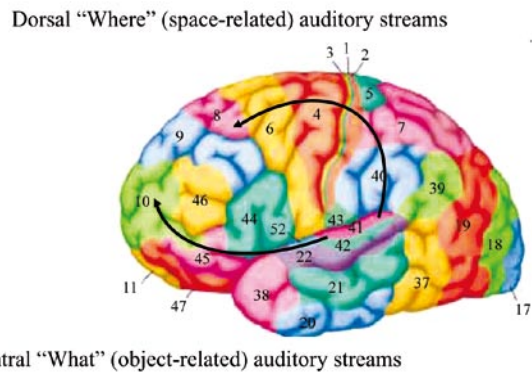


Fig. 3 Prefrontal connections of auditory cortex in macaque monkeys. Adapted from reference^[19].

nonspatial (i.e., areas 10, 12vl, rostral 46d) areas in prefrontal cortex^[16, 17, 19, 27, 28] (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^[29].

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)^[30-34] (for a review see^[18]; 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^[35]. 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;

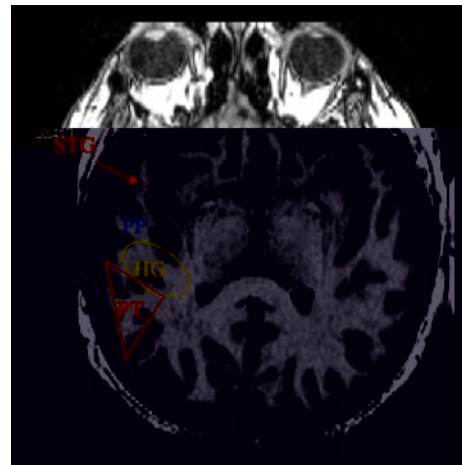


Fig. 4 Schematic drawings of the horizontal view of principal auditory areas in the human left STG. Adapted from reference^[34].

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^[36]. 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.*^[37], and the early and later stages used in study of Arnott *et al.*^[38]. 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^[35,38] and using speech samples^[37].

Maeder *et al.*^[39] 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 spatialized than to meaningful sounds^[39], which indicated that the differential activation of the two pathways resulting from an organizational principle of human auditory cortex rather than the possible

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^[45,46]. 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^[47] 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^[47].

Similar to the visual system, the interaction between the

auditory "where" and "what" pathways reflects the brain

s r (y -) 6

auditory stimuli for performing either a semantic content recognition or a spatial position localization matching task^[51]. 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^[52,53].

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-

The auditory dual-pathway in human cortex

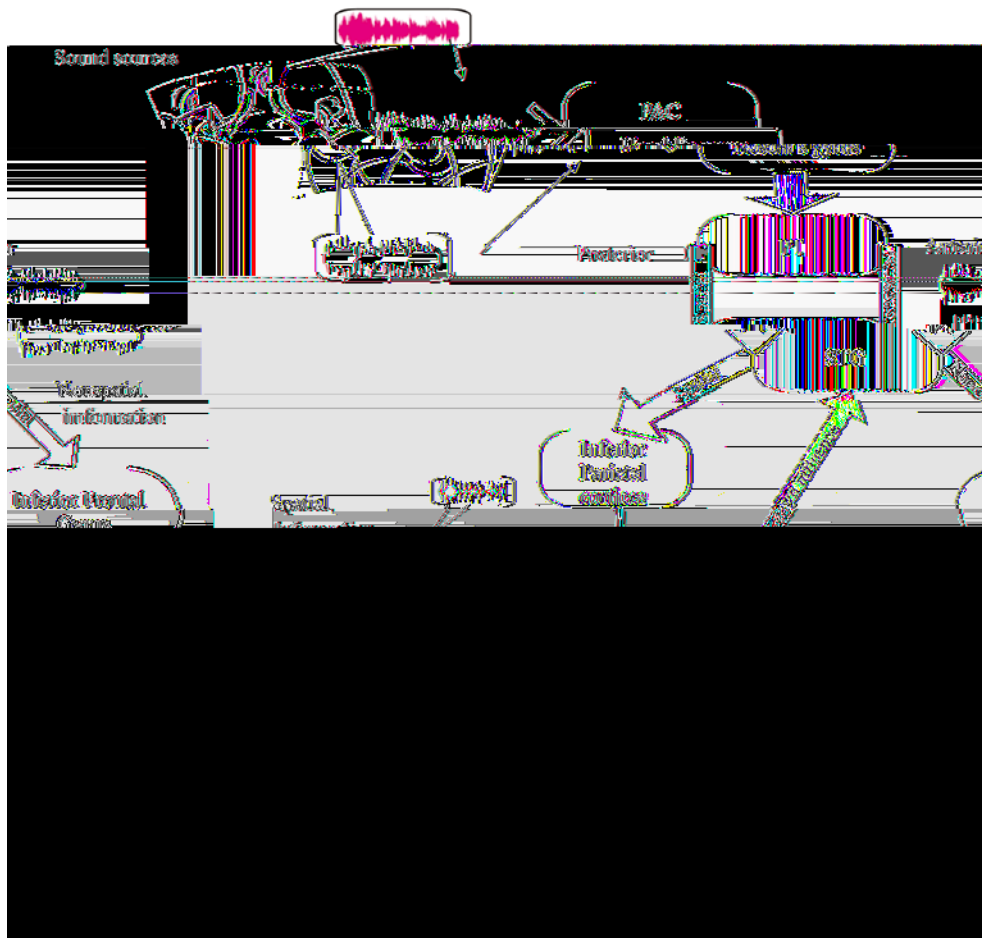


Fig. 5 The dual-pathway model for explaining auditory processing in human cortex. The wave color of red and blue indicate the different pitch information. The yellow pathway indicates the spatial information pathway. The pink pathway indicates the nonspatial information pathway.

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^[54]. 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^[55,56]. Indeed, when the auditory task requires listeners to selectively compare or evaluate the location of sound source, IPL activity is specifically enhanced^[54,57,58], 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^[59-61], allowing the brain to further encode target-signal details (for a meta-analysis see^[62]). 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^[45,46].

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^[63]. 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 Neurootol* 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