Review

Tonotopic mapping of human auditory cortex

Melissa Saenz a,b,*, Dave R.M. Langers c

aLaboratoire de Recherche en Neuroimagerie (LUREN), CHUV, Department of Clinical Neurosciences, Lausanne University Hospital, Mont Puisilble 16, Lausanne 1011, Switzerland
bInstitute of Bioengineering, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015, Switzerland
cNational Institute for Health Research (NIHR) Nottingham Hearing Biomedical Research Unit, School of Clinical Sciences, University of Nottingham, Queen’s Medical Centre, Nottingham, United Kingdom

Article info

Article history:
Received 7 May 2013
Received in revised form 19 July 2013
Accepted 25 July 2013
Available online xxx

Abstract

Since the early days of functional magnetic resonance imaging (fMRI), retinotopic mapping emerged as a powerful and widely-accepted tool, allowing the identification of individual visual cortical fields and furthering the study of visual processing. In contrast, tonotopic mapping in auditory cortex proved more challenging primarily because of the smaller size of auditory cortical fields. The spatial resolution capabilities of fMRI have since advanced, and recent reports from our labs and several others demonstrate the reliability of tonotopic mapping in human auditory cortex. Here we review the wide range of stimulus procedures and analysis methods that have been used to successfully map tonotopy in human auditory cortex. We point out that recent studies provide a remarkably consistent view of human tonotopic organisation, although the interpretation of the maps continues to vary. In particular, there remains controversy over the exact orientation of the primary gradients with respect to Heschl’s gyrus, which leads to different predictions about the location of human A1, R, and surrounding fields. We discuss the development of this debate and argue that literature is converging towards an interpretation that core fields A1 and R fold across the rostral and caudal banks of Heschl’s gyrus, with tonotopic gradients laid out in a distinctive V-shaped manner. This suggests an organisation that is largely homologous with non-human primates.

This article is part of a Special Issue entitled ‘Human Auditory Neuroimaging’.

© 2013 Published by Elsevier B.V.

1. Introduction

Neuroimaging techniques like electro- and magnetoencephalography (EEG, MEG) or positron emission tomography and functional magnetic resonance imaging (PET, fMRI) are painting an increasingly detailed picture about how the human brain is organised. Numerous brain networks have been identified that consistently show coherent patterns of activity during a variety of tasks, and even during rest (Fox et al., 2005; Gazzaniga, 1989). The sensory modalities provide excellent examples of brain networks for which parcellation into subdivisions has been achieved. That is because several sensory systems feature faithful representations of the peripheral sensory epithelia (Kaas, 1997; Weinberg, 1997). For example, neurons in the visual cortex are tuned to particular areas in the visual field of view, and are laid out on the surface of the cerebral cortex in a fashion that can be mapped one-on-one onto the extent of the retina. In fact, multiple such topographic representations exist, each within a distinct subdivision of the visual cortex (Engel et al., 1997; Wandell and Winawer, 2011). Similarly, the somatosensory cortex that is involved in the sense of touch features representations of the various body parts, giving rise to the cortical homunculus. Again, multiple such somatotopic maps appear in parallel in several adjacent cortical subdivisions, and a similar map occurs in neighbouring motor cortex (Mattay and Weinberger, 1999; Narici et al., 1991; Sanchez-Panchuelo et al., 2012).

The auditory system receives input from the organ of Corti in the inner ear. Hair cells are laid out along the length of the basilar membrane, spiralling along the windings of the cochlea. The nerve fibres that synapse with the hair cells retain this essentially one-dimensional cochleotopic organisation, all the way up to the auditory cortex. Because the mechanical properties of the basilar membrane gradually change along its length, hair cells are tuned to progressively higher frequencies when traversing the cochlea from its apex to its base. Thus, the inner ear acts as a sound frequency analyser, and sound information is transmitted centrally along...
numerous frequency channels in parallel (Ellis and Helmholtz, 1885; Fuchs, 2010; Meyer and Moser, 2010; Von Bekesy, 1949). Due to this frequency-place code, cochleotopy is more commonly referred to as tonotopy (\(\text{tόνωτος} = \text{tone}; \text{tόπος} = \text{place}\)) in the neuroimaging literature.

The ability to determine tonotopic maps not only serves to gain insight in the functional organisation of the auditory system regarding frequency processing, which may be argued to be one of the most basic functions it performs. Besides that, it provides a tool to parcelle the central auditory system into meaningful subdivisions of which the distinct properties can be studied with regard to acoustic features other than frequency, as well as non-acoustic factors like attention. Multiple tonotopic progressions can be found in various subdivisions of the auditory nuclei in the brainstem, midbrain, and thalamus, and in the auditory cortex of the cerebrum, as illustrated in Fig. 1 (Clopton et al., 1974; Günter Ehret and Romand, 1997; Rees and Palmer, 2010). Currently, frequency is the only acoustic parameter that is unequivocally held to be topographically mapped, although other parameters like sound intensity (Bliecdn et al., 2002; Pantev et al., 1989a), tuning bandwidth (Moerel et al., 2012; Seifritz et al., 2006), and modulation rate (Langner et al., 1997; Barton et al., 2012; Herdener et al., 2013) have been suggested to form complementary maps.

Tonotopic mapping of the auditory cortex has proven particularly challenging for human neuroimaging. This is in part due to the small size of auditory cortical fields relative to the spatial resolution of neuroimaging techniques, and in part due to a lack of consensus regarding architectonic definitions of human primary auditory cortex. As a result, neuroimaging studies of human tonotopy have proposed different, even opposing, views regarding the orientation of the primary tonotopic gradients in auditory cortex, which in turn leads to different predictions about the locations of specific auditory fields. However, despite differences in map interpretation and a variety of experimental paradigms, we emphasize that virtually all recent studies show a remarkably consistent spatial pattern of frequency preference in human auditory cortex. Here, we review the history of neuroimaging of tonotopy, critically review the differing map interpretations, and describe the range of experimental paradigms used thus far.

2. Tonotopic organisation in humans

2.1. Extrapolating from animal studies

By means of invasive animal studies, the existence of tonotopic progressions has been shown for many subdivisions in the central auditory system (Clopton et al., 1974; Ehret and Romand, 1997; Rees and Palmer, 2010). All subdivisions in the brainstem are tonotopically organised (Fig. 1): in the cochlear nucleus, tonotopic progressions exist in the anteroventral subdivision as well as in the dorsal and neighbouring posteroverentral divisions, while in the superior olivary complex tonotopic organisations have been reported for both the medial and lateral subdivisions as well as the medial nucleus of the trapezoid body (Kandler et al., 2009; Ryugo and Parks, 2003). In the midbrain, two pathways diverge (Hu, 2003; Müller and Rollins, 2002). One is the lemniscal classical auditory pathway that is tonotopically organised throughout. It comprises the central nucleus of the inferior colliculus and the ventral division of the medial geniculate body, which project to primary areas in auditory cortex. The other is the extralemniscal non-classical auditory pathway that shows a diffuse frequency organisation and provides aspecific sensory information. It comprises the cortex of the inferior colliculus and the dorsal and magnocellular subdivisions of the medial geniculate body, and projects to non-primary auditory cortex as well as various non-auditory brain structures involved in multimodal, associative, and affective processing.

The organisation of the auditory cortex on the superior temporal gyrus of the cerebrum has been most extensively studied. Tono- topic progressions were observed in numerous animal species, including birds (Capsius and Leppelsack, 1999; Cohen and Knudsen, 1996; Terleph et al., 2006), rodents (Hellweg et al., 1977; Kelly et al., 1986; McMullen and Glaser, 1982; Merzenich et al., 1976; Stiebler et al., 1997), primates (Kusmierek and Rauschecker, 2009), including birds (Capsius and Leppelsack, 1999; Cohen and Knudsen, 1996; Terleph et al., 2006), rodents (Hellweg et al., 1977; Kelly et al., 1986; McMullen and Glaser, 1982; Merzenich et al., 1976; Stiebler et al., 1997), primates (Kusmierek and Rauschecker, 2009; Luethtke et al., 1989; Morel and Kaas, 1992; Scott et al., 2011), and other mammals (Reale and Imig, 1980; Suga and Jen, 1976). In non-human primates, a hierarchical model of auditory cortical organi- sation has emerged based on combined knowledge of electrophysiology, cortical architecture, and connectivity. In this model, an elongated core (primary regions) is comprised of up to three roughly collinear tonotopic fields (primary auditory field A1, followed by a rostral field R and an even more rostral temporal field RT) surrounded by several belt fields (secondary regions), further surrounded by higher-order parabelt fields (Kaas and Hackett, 2000).

The elongated core is situated along a posterior-to-anterior axis. Along this axis, neuronal frequency preferences follow a gradient of high to low (A1), followed by a reversed gradient of low back to high (R), followed by a third smaller and perhaps less clearly organised gradient of high back to low (RT). Thus, the overall pattern is a clear “high-to-low-to-high” corresponding to core fields A1 and R followed by a less distinct “high-to-low” corresponding to core-like field RT, with the borders between fields marked by frequency gradient reversals. Architectonic and histo-chemical markers of the primary core are most pronounced and similar in A1 and R and somewhat less distinct in RT (Imig et al.,

Fig. 1. The central auditory pathway. All nuclei that form part of the classical lemniscal auditory pathway are tonotopically organised. These include various subdivisions of the cochlear nucleus (CN), superior olivary complex (SOC), inferior colliculus (IC), and medial geniculate nucleus (MGN). In the auditory cerebral cortex in the superior part of the temporal lobe, expected divisions of core, belt, and parabelt are based on the non-human primate model of auditory cortical organisation. Human neuroimaging consistently shows at least two primary tonotopic gradients (“high-to-low-to-high”) in the auditory cortex, homologous to primary fields A1 and rostral field R in the monkey cortex. In some primate studies, a third rostrottemporal field RT is delineated, but neuroimaging evidence for a similar field in humans is sparse.

Please cite this article in press as: Saenz, M., Langers, D.R.M., Tonotopic mapping of human auditory cortex, Hearing Research (2013), http://dx.doi.org/10.1016/j.heares.2013.07.016
1977; Morel et al., 1993; Kaas and Hackett, 2000). These primary
tonotopic gradients extend seamlessly into a number of the
neighbouring belt areas (Kusmierek and Rauschecker, 2009; Morel
et al., 1993; Petkov et al., 2006).

As pointed out in a recent review by Baumann et al. (2013), the
primate tonotopic progressions from posterior to anterior are not
strictly collinear, but rather follow a distinctly angled pattern as has
been observed in macaques and marmosets (Morel et al., 1993;
Kosaki et al., 1997; Kaas and Hackett, 2000; Bendor and Wang,
2008; Baumann et al., 2010). The A1 gradient is angled from a
more-medial starting point to a more-lateral end point, and the R
gradient is angled from the more-lateral starting point to a more-
medial end point. Thus, the primary gradients of A1 and R (“high-
to-low-to-high”) form a V-shape, with the low frequency mid-zone
positioned more laterally and the two high frequency end points
positioned more medially. Another key point raised by Baumann
et al. (2013) is that the macaque temporal plane is not
fl

at, as
often assumed. Rather, there is an often-overlooked protuberance
at the location where the posterior auditory cortex turns downward
towards the anterior auditory cortex. In some cases this protuber-
ance is pronounced enough to form a mini-gyrus (the annexant
gyrus) which resembles a rudimentary Heschl’s gyrus (HG) (Jones
et al., 1995). This protuberance predicts very consistently the low
frequency area at the border of A1 and R (Baumann et al., 2010)
with tonotopic gradients running across it. Both of these observa-
tions prove to be informative when comparing to the layout of
human tonotopic gradients, as discussed below.

2.2. Architectonic parcellations in humans

After more than a century of mapping human cortical archi-
tecture, a complete model of human auditory cortical organisation
remains elusive. Early architectonic studies identified a bilateral
region on the temporal plane with the characteristics of primary
sensory cortex including a well-developed granular layer 4
(koniocortex), dense myelination, and thalamic connectivity
(Fleschig, 1908; Campbell, 1905; Brodmann, 1909; von Economo
and Koskinas, 1925; von Economo and Horn, 1930). This region is
usually referred to as primary auditory cortex (PAC) in the human
literature and shares many architectonic features with the auditory
core in non-human primates (Hackett et al., 2001). Across studies,
PAC co-localises approximately with the medial two-thirds of HG,
but the gyral borders do not reveal the exact architectonic borders
(Fig. 2a). PAC has been noted in some cases to reach anteriorly onto

Fig. 2. Tonotopic Map Layout and Interpretations. Panels show a view of a partially inflated cerebral surface, looking into the Sylvian fissure that separates the temporal lobe below from the frontal and parietal lobes above; light grey: gyral convexities; dark grey: sulcal concavities. (a) Primary auditory cortex colocalises with Heschl’s gyrus (HG), a transverse superior temporal gyrus, separating the planum polare (PP) on the anterior side from the planum temporale (PT) posteriorly. (b) Interpretations of the tonotopic organisation in humans have historically developed from a single effective gradient spanning HG, to a pair of oppositely collinear progressions stretching along HG, to a pair of oblique V-shaped progressions folding across HG. See main text Section 2 for an elaborate discussion of these developments. (c) Single-subject and mean group-level (N = 40) best-frequency maps illustrate typically observed tonotopic patterns. Low frequencies are found on the crest of HG flanked by two high-frequency zones posteromedially towards the planum temporale and anteromedially towards the circular sulcus. This gives rise to two cortical subregions with systematic tonotopic progressions. The right panel shows only these two regions, located on the rostral and caudal banks of HG, in which tonotopic gradient vectors achieve consistent orientation and non-zero magnitude across subjects. These regions are thought to correspond with human homologous fields A1 and R, folding across HG with gradients oriented in a distinct V-shaped pattern. Data from Langers (2013); acquired at 3 T.

Please cite this article in press as: Saenz, M., Langers, D.R.M., Tonotopic mapping of human auditory cortex, Hearing Research (2013), http://dx.doi.org/10.1016/j.heares.2013.07.016
the planum polare (PP) and posteriorly onto the planum temporale (PT) (Brodmann, 1909; von Economo and Koskinas, 1925; von Economo and Horn, 1930; Rademacher et al., 1993; Morosan et al., 2001). Early studies (von Economo and Koskinas, 1925; von Economo and Horn, 1930) identified a relatively broad koniocortical area not fully contained within HG and noted that its center, located along the crest of the gyrus, displayed the clearest koniocortical features. Later studies identified finer-scale architectonic and histochemical inhomogeneities and proposed different sub-parcellations of PAC (Galaburda et al., 1978; Rivier and Clarke, 1997; Morosan et al., 2001; Wallace et al., 2002; Sweet et al., 2005; Fullerton and Pandya, 2007). The different parcellations vary in the number and location of identified fields (for a review, see Clarke and Morosan, 2012).

Thus, it appears that human PAC is architectonically heterogeneous and while there is general agreement on the location of its centre, there is a lack of agreement on its exact areal borders and number of subdivisions. High inter-subject and inter-hemispheric anatomical variability is a complicating factor, including commonly forked or duplicated HG (estimated occurrence 41%, Rademacher et al., 1993). Nevertheless, most studies do identify an elongated posteromedial-to-anterolateral densest core region along the crest of HG that appears similar in shape to the elongated posterior-to-anterior auditory core in the monkey (Hackett et al., 2001), although distinctly rotated. This led to the expectation that the human auditory core is rotated compared to monkeys and that the homologues of A1, R, and RT and their tonotopic gradients would be found to run along the length of HG. On the other hand, if no rotation between human and monkey is presumed, then the primary tonotopic gradients would be expected to traverse a posterior-to-anterior direction roughly across HG, rather than along it. Hence, given the ambiguities in human architectonic definitions, there are reasonable arguments to expect that tonotopic gradients could occur along either of two opposing orientations.

2.3. An effective tonotopic progression along Heschl's gyrus

The earliest non-invasive studies that suggested the existence of a tonotopic organisation in healthy humans employed MEG (Elberling et al., 1982; Romani et al., 1982). These initial observations were later confirmed by a large number of other MEG studies, reviewed by Pantev and Lütkenhöner (2000), as well as several EEG and chronic microelectrode studies (Cansino et al., 1994; Fujioka et al., 2002; Gabriel et al., 2004; Hoke et al., 1998; Howard et al., 1996; Huotilainen et al., 1995; Kuriki and Murase, 1989; Langner et al., 1997; Lütkenhöner and Steinsträter, 1998; Pantev et al., 1988, 1994; Tiitinen et al., 1993; Verkindt et al., 1995; Weisz et al., 2004; Yamamoto et al., 1988)). Overall, the activation dipole's depth below the scalp and its coordinate along the rostrocaudal axis were found to increase with stimulus frequency, and its orientation varied due to gyral morphology. These findings suggested that an effective frequency progression extends along HG, with a low-frequency starting point at the lateral side of HG and a high-frequency end point at its medial side (Fig. 2b, left panel). A number of authors attempted to determine whether this topographic organisation is more closely related to the frequency (i.e. spectral content) or the pitch (i.e. fundamental frequency) of the stimulus, but results were contradictory and inconclusive (Cansino et al., 2003; Crottaz-Herberte and Ragot, 2000; Diesch and Luce, 1997; Fujioka et al., 2003; Pantev et al., 1998b).

Dipole locations summarise the average location of neural activity of a spatially extended region of cortex. Using laborious experimentation, MEG may distinguish the relative positions of dipoles that result from different sound stimuli in a single individual even if they differ by only a few millimetres (Lütkenhöner and Steinsträter, 1998). Moreover, by separately reconstructing the dipoles that correspond with the various resolvable peaks in the time signal, which are generated at hierarchically distinct processing levels, multiple tonotopic maps may be resolved even if they are spatially proximate (Pantev et al., 1995; Verkindt et al., 1995). This may in principle allow MEG/EEG to assess even a complex tonotopic layout that consists of abutting progressions in more than one cortical subdivision. However, nearby dipoles that occur simultaneously remain impossible to resolve, and an accurate absolute localisation depends on a proper dielectric model of the head. In practice, the dynamic movement of dipoles across the cortex further complicates the interpretation of functional outcomes (Ozaki and Hashimoto, 2007). This may explain why according to MEG tonotopic progressions are often absent or disorderly, and highly variable across subjects (Lütkenhöner et al., 2003). Although MEG and EEG provide excellent temporal resolution, alternative neuroimaging techniques that offer more precise spatial resolution and the ability to sample large numbers of nearby cortical sites simultaneously are therefore preferable.

Early PET and fMRI studies determined the effective location of sound-evoked activation in each hemisphere (e.g., the centre of mass of an activation cluster, or the location of its peak activation) in response to as little as two different tone frequencies, but still confirmed that the higher frequency was represented more postero/medially along HG than the lower frequency (Bilecen et al., 1998; Lauter et al., 1985; Lockwood et al., 1999; Wassinger et al., 1997). Later studies included more frequencies in order to show that the tonotopic progression in PAC was gradual (Langers et al., 2007; Le et al., 2001; Ottaviani et al., 1997; Petkov et al., 2004; Scarf et al., 2004; Yetkin et al., 2004). These PET/fMRI studies therefore well agreed with the MEG/EEG literature. At that time, this resulted in the view that a large portion along the extent of HG corresponded with one core auditory field.

2.4. Multiple tonotopic maps per hemisphere

Given the existence of multiple core fields in numerous animal species, several researchers subsequently endeavoured to discover the “missing” other maps in humans by exploiting the spatial resolution of fMRI in order to distinguish multiple tonotopic gradients per hemisphere. Talavage et al. (2000) identified eight consistently occurring response foci to either low- or high-frequency stimuli. The existence of a largely consistent set of foci was supported in a subsequent study by Schönwiesener et al. (2002). However, the latter authors raised the question whether regions tuned to low- or high frequencies should be regarded as evidence for a tonotopic organisation containing gradual frequency progressions or, alternatively, whether different cortical subareas just happen to rely more heavily on low- or high-frequency content for the functions that they perform. This issue was addressed in a follow-up study by Talavage et al. (2004), who showed that these foci were pairwise connected by six tonotopic gradients on the basis of waves of activation that travelled across the cortex in response to slow frequency sweeps. Multiple low- and high-frequency gradients were oriented around the lateral-to-medial direction, consistent with the effective progression that was known to exist. But results also included an oppositely oriented gradient on the lateral side that did not fit into the existing picture. This work was the first to hint at a much more complicated tonotopic organisation in humans that also comprises non-primary auditory subdivisions.

In the same period, a reversed tonotopic organisation in lateral temporal cortex, with low-frequency responses occurring in slightly more medial locations than high-frequency responses, was reported by Yang et al. (2000). The location of this gradient fitted...
into the idea that, like in primates, two abutting frequency gradients might exist, the one known to extend more or less along the medial half of HG, and another one positioned in an adjacent more lateral location. In follow-up work, evidence for both gradients was shown (Engelien et al., 2002). Through the use of a high-field strength scanner, highly detailed and much more convincing evidence soon appeared for the simultaneous existence of two mirror-symmetric tonotopic maps in adjacent subdivisions of PAC (Formisano et al., 2003). These maps were oppositely directed and extended more or less collinearly along the axis of HG, touching at their low-frequency boundary (Fig. 2b, middle panel). This lent support for the interpretation that HG hosts the core auditory cortex in humans, featuring a high-to-low-to-high frequency representation.

In the second half of that decade, comparable tonotopic maps were revealed using lower field strengths (Hertz and Amedi, 2010; Riecke et al., 2007; Seifritz et al., 2006; Upadhyay et al., 2007; Woods et al., 2009). These were argued to be consistent with the existing evidence that showed two primary cortical subdivisions. As reviewed by Woods and Alain (2009), these likely form the human homologues of the fields A1 and R in non-human primates.

2.5. Along or across Heschl’s gyrus?

At the start of this decade, Humphries et al. (2010) proposed an alternative tonotopic organisation. Although the high-to-low-to-high frequency representation remained intact, these authors argued that the tonotopic progressions run perpendicularly across HG, rather than parallel along HG. According to their data, an elongated zone on HG was found to respond preferentially to lower frequencies, whereas zones posterior and anterior to HG were more sensitive to higher frequencies. This view was subsequently subscribed to and elaborated on by various other groups (Striem-Amit et al., 2011; Da Costa et al., 2011; Langers and van Dijk, 2012; Herdener et al., 2013).

Showing high-resolution individual subject mappings, Da Costa et al. (2011) reported that the primary tonotopic gradients consistently ran across HG, and pointed out that the low frequency border between the two gradients of A1 and R was consistently centered on the full HG, regardless of individual morphological variations (i.e. whether single, forked, or duplicated). This functional-anatomical relationship is reminiscent of the macaque auditory cortex where an anatomical protuberance very consistently predicts the low frequency border between A1 and R (Baumann et al., 2013). Their findings significantly revised HG as a marker for human PAC, since previously it was commonly assumed that PAC would occupy only the anterior part of forked or duplicated HG (for a review, see Abdul-Kareem and Sluming, 2008).

2.5. Along or across Heschl’s gyrus?

On the basis of group-level data, Langers and colleagues similarly found gradients running across HG with fields A1 and R on the caudal and rostral flanks of HG, respectively (Langers and van Dijk, 2012; Langers et al., 2012). These authors specifically noted that the direction of these tonotopic gradients pointed diagonally, resulting in an angled V-shaped pair of frequency progressions that is consistent with the anterior-to-posterior V-shaped tonotopic axis in primates (Kaas and Hackett, 2000). This observation also explains why previous studies with lower spatial resolution, unable to differentiate between the two medial high-frequency end points, suggest only a single lateral-to-medial tonotopic progression effectively. There is variability across individuals in the exact angle of the V-shape, and the two high-frequency end points are especially close together in individuals with more sharply angled gradients.

The interpretation of V-shaped gradients folding across HG (Fig. 2b, right panel) is harmonious with the human non-primate model in terms of axis orientation and does not require the presumption of a significant reorientation during human evolution. As pointed out by Baumann et al. (2013), this model is functionally and anatomically parsimonious with the macaque auditory core: considering that in the macaque angled gradients run across the above mentioned low-frequency protuberance which may be an anatomical precursor to HG. This interpretation remains somewhat controversial, however. In a pair of recent publications, interpretations were made that remain consistent with the along-HG view of Formisano et al. (2003). However, still more recently, papers appeared that directly challenged those views and remain consistent with the across-HG view of Humphries et al. (2010).

Moerel et al. (2012) concluded that in a region that more or less coincided with the axis of HG a narrower frequency tuning occurs than in surrounding cortical areas. When projecting the tonotopic gradient direction onto this axis, direction reversals were observed. However, it is a strong assumption that tonotopic gradients must align along an axis of narrow tuning. Langers (2013) performed a statistical assessment of local gradient direction and magnitude that did not rely on any predefined axis of interest, and results revealed two consistent gradients running across HG with A1 and R extend along HG. However, on the basis of an experiment with a comparable design, Herdener et al. (2013) suggested an entirely different, larger-scale periodotopic organisation. Their findings are in line with primary tonotopic gradients running across Heschl’s gyrus.

Dick et al. (2012) and Lutti et al. (2013) employed non-invasive T1-mapping to localise putative core auditory areas, and found an elongated keyhole-shaped region on the medial two-thirds of HG. Within this putative core region, they find a posteromedial-to-anterolateral progression from high frequencies to a low-frequency trough consistent with the expected A1 gradient. This was followed by a reversed low-to-medium frequency progression directed anteromedially from the low-frequency trough. Their findings are consistent with angled gradients in an overall posterior to anterior orientation, but their conclusions are more nuanced by questioning whether a complete reversal to high-frequencies is found within the core.

Even if the interpretation of tonotopic maps remains an issue of some debate, it should be pointed out that the maps of cortical frequency representation show remarkable consistency across recent studies. Virtually all of them show a clear low-frequency representation on the mid-to-lateral half of HG. This low-frequency “trough” is flanked by high-frequency representations rostromedially towards the PP and caudomedially towards the PT. Representative functional outcomes are illustrated in Fig. 2c. In retrospect, even the results of studies that originally gave rise to different interpretations tend to show these features. Currently, a dozen studies support essentially the same tonotopic map.

In our view, it is because the low-frequency zone between A1 and R is a trough that different gradient orientations can be interpreted around it. If tonotopic maps are considered on their own, the interpretation that gradients “high-to-low-to-high” traverse across Heschl’s gyrus is most apparent. On the other hand, if (based on combination with other data) core tonotopic gradients are
presumed to run along the axis of HG, then an initial “high-to-low” gradient may be interpreted followed by a partial reversal on the lateral extent of HG. However, on lateral HG it is difficult to account for the return to high-frequencies that would be expected given tonotopic mappings in primates (Bendor and Wang, 2008; Baumann et al., 2010; Tanji et al., 2010).

In our opinion, based on the latest findings discussed above, it is most plausible that core fields A1 and R in humans are positioned on the caudal and rostral banks of HG, respectively. This makes them extend across HG, rather than along. At the same time, tonotopic gradients are set in an oblique V-shaped orientation. This suggests an organisation that is largely homologous with non-human primates.

2.6. Beyond primary auditory cortex

Besides the two tonotopic progressions in PAC, evidence for additional maps in other subdivisions has also been reported, although most of these findings currently remain to be confirmed.

At a cortical level, an additional low-frequency focus has been observed posterior to HG on the lateral PT on the superior temporal gyrus (Humphries et al., 2010; Langers and van Dijk, 2012; Talavage et al., 2004). This may connect to the posterior high-frequency end points, medial to HG, extending the primary tonotopic organisation. Thus, together with the primary end points, a zig-zag high-to-low-to-high-to-low frequency progression is obtained. This alternation of low and high frequency preferences is reminiscent of the repeated reversals that were observed by means of high-resolution fMRI in primate cortex (Pekov et al., 2009; Tanji et al., 2010). Additional tonotopic gradients may exist in the extreme lateral part of the superior temporal gyrus, neighbouring the superior temporal sulcus and middle temporal gyrus (Striem-Amit et al., 2011). These areas are difficult to assess because they are poorly responsive to tones, but analysis techniques have been proposed that allow the usage of paradigms that employ broadband stimulation (De Martino et al., 2013; Moerel et al., 2012).

At a subcortical level, differential frequency-dependent responses have been reported in the human medial geniculate body by means of intracerebral EEG recordings (Yvert et al., 2002). More recently, evidence for a tonotopic organisation in the human inferior colliculus was shown using fMRI (De Martino et al., 2013). Dorsolaterally, lower frequencies were found to be represented than in deeper ventromedial locations, consistent with findings in humans obtained with electrical stimulation (Lim et al., 2013) and with imaging outcomes from the inferior colliculus in rats (Cheung et al., 2012). The neuroimaging of subcortical auditory nuclei, and the lower brainstem nuclei in particular, remains an enormous challenge, however.

2.7. Tonotopic reorganisation

Tonotopic mapping by means of neuroimaging has already found application in relation to the study of plastic reorganisation in auditory cortex. This may occur at a short term in the normal brain, or at longer terms in relation to various chronic auditory as well as non-auditory disorders.

Tonotopic representations may undergo experience-dependent plasticity, for instance by means of classical conditioning (Morris et al., 1998) or frequency-specific sound deprivation (Pantev et al., 1999). Following hearing loss, the apparent extent of cortical tonotopic organisation was found to have shrunk in otosclerotic patients, but gradually recovered over the course of a few weeks after corrective surgery involving stapes substitution (De Campora et al., 2003; Tecchio et al., 2000). It has similarly been suggested that tonotopic organisation may be measured in patients with cochlear implant devices by stimulating different electrode channels (Ponton et al., 1993; Thai-Van et al., 2010), but in practice it has proven difficult to distinguish gradual tonotopic progressions from responses arising from distinct cortical fields (Guiraud et al., 2007; Seghier et al., 2005). Finally, tonotopic abnormalities play a role in pathophysiological models of tinnitus and hearing loss. Such abnormalities were initially reported using MEG (Mühlner et al., 1998; Wienbruch et al., 2006), but these findings could later not be confirmed by means of high-resolution fMRI (Langers et al., 2012; Langers, 2013). With regard to non-auditory disorders, the tonotopic organisation of the auditory cortex was reported to be disturbed in schizophrenia (Rojas et al., 2002) and to be expanded in the blind (Elbert et al., 2002), although the latter was not observed in a later study (Stevens and Weaver, 2009).

Overall, evidence for tonotopic reorganisation in humans is weak. On the basis of animal research, it is plausible that abnormal tonotopic maps arise in various conditions. Most of the above neuroimaging studies offered poor spatial resolution however, especially when compared to the tonotopic mapping studies that have appeared in recent years. Furthermore, results were often inconsistent. Still, given the emerging consensus and ever more detailed insight in cortical organisation, tonotopic mapping offers exciting possibilities for practical applications to be exploited in coming years.

3. Paradigms

As reviewed above, virtually all recent high-resolution fMRI studies converge upon a remarkably similar spatial layout of frequency preferences in human PAC (Fig. 1). In this section, we review the variety of stimulus types and data analysis techniques successfully employed thus far. The convergent results across experimental paradigms suggest that the underlying tonotopic signals are highly robust.

3.1. Stimulus frequency and intensity level

A first step in designing a mapping experiment is to choose a range of stimulation frequencies that adequately sample the map. In humans, the basilar membrane is stimulated at different locations along its length as a logarithmic function (approximate equal-octave spacing) of sound frequencies ranging from roughly 20 to 20,000 Hz, and that spacing is maintained in central auditory maps to the cortex (Merzenich et al., 1975). Sensitivity is highest in the mid-frequency range (discussed further below) and a gradual loss of high-frequency hearing begins early in normal ageing. Most tonotopic mapping studies have thus employed logarithmically-spaced tonal stimuli avoiding low and high frequency extremes (for example: 250, 500, 1000, 2000, 4000, and 8000 Hz).

Characteristic frequency (CF) is defined as that pure tone frequency at which a neuron achieves its most sensitive threshold. In turn, the threshold equals the sound intensity that is required to elevate the activity of the neuron to a certain level above its spontaneous activity. This definition therefore requires a tone's intensity to be varied to determine thresholds, and its frequency to be varied to determine CFs. This gives rise to a large number of intensity—frequency combinations for which responses need to be measured (Recanzone et al., 2000). For non-invasive neuroimaging methods, averaging over many comparatively slow measurements is needed, making this an impractical approach. For that reason, a best frequency (BF) is simply defined as that frequency that evokes the strongest response. Thus, a tone's intensity can remain fixed, and only its frequency needs to be varied. Although this BF measure must converge to the CF at sufficiently low sound intensities, tuning curves often show asymmetric widening towards higher intensities,
and complicated non-monotonic or bimodal response characteristics may further complicate the relationship (Sutter, 2000).

In fMRI, sound presentation levels are relatively high (typically 50–80 dB HL), turning this into an issue. Furthermore, fMRI does not have the spatial resolution to measure responses of individual neurons, but at best measures the response of hundreds of thousands of neurons collectively. Both the intensity-related response widening and the agglomeration across populations of neurons explain why frequency preferences as measured using neuro-imaging are much broader than the often narrow tuning that is observed for individual neurons using invasive electrophysiology techniques. This leads to the question whether frequency preferences can at all be reliably measured with fMRI.

By averaging out the intricate response properties of individual neurons, a response correlate should be achieved that captures the relative frequency tuning, in the sense that sites with higher BF contain neurons with higher CFs on average. The clearest confirmation of this assumption comes from high-field fMRI studies of tonotopic organisation in the macaque (Petkov et al., 2006; 2009; Tanji et al., 2010; Baumann et al., 2010). These studies employed supra-threshold tone stimuli (70–90 dB) and revealed frequency gradients that matched the expected size and locations of A1, R, and RT, as known from previous invasive studies. Other fMRI mapping studies have directly compared outcomes obtained with softer versus louder mapping stimuli (Woods et al., 2009; Tanji et al., 2010; Langers and van Dijk, 2012) and found that blood oxygenation level-dependent (BOLD) response amplitudes increase with stronger stimuli but overall patterns of frequency preference remain the same.

It is also important to note that perceived loudness varies across frequencies with highest sensitivity in the 2–5 kHz range, mostly due to the transfer function of the ossicles of the inner ear. For example, a 2-kHz tone results in more cochlear activation and higher perceived loudness than a 200-Hz tone of the same intensity level. Tonotopic mapping experiments have taken various steps to better equate stimulation levels across frequencies. One approach is to adjust sound intensity levels to a fixed level (e.g. 40 dB) above subjects’ behaviourally measured hearing thresholds, per frequency (Talavage et al., 2004; Langers and van Dijk, 2012). A second approach is to adjust intensity levels according to standard equal loudness curves (equal phon, ISO 226) (Da Costa et al., 2011, 2013). Another reasonable option would be to let subjects behaviourally match the loudness of each test frequency used to that of a reference frequency of a fixed intensity level.

3.2. Sparse vs. continuous imaging

To lessen the impact of acoustic scanner noise (typically exceeding 100 dB without the use of attenuation), sparse or clustered imaging protocols have been developed in which silent gaps are inserted between successive scans or clusters of scans, as opposed to standard continuous imaging (Edmister et al., 1999; Hall et al., 1999). Experimental stimuli are presented during the silent gaps and, due to haemodynamic delay, the stimulus response is measured by resuming scanning after stimulus offset. Both sparse and continuous imaging protocols have been successfully employed in the imaging of tonotopic maps. Scarf et al. (2004) showed that sound stimuli near the most powerful frequencies in the acoustic spectrum of the scanner noise were partially masked in loudness, and cortical activation at those frequencies was lowest. By changing the parameters of the acquisition sequence, they shifted the dominant frequency of the scanner noise and showed that the observed activation minima shifted along with that. Their results suggest that sound frequencies that are present in the scanner noise will tend to be underrepresented in the obtained tonotopic maps. This may be caused by a neural or haemodynamic saturation in cortical sites that are tuned to the ongoing scanner noise. Interestingly, Woods and Alain (2009) directly compared maps obtained with sparse and continuous protocols and reported essentially the reverse: mid frequencies that were closest to the peak scanner noise evoked the strongest responses. They attributed this to long-term potentiation or to an adaptation of inhibitory input. In their study, response magnitudes were larger in the sparse imaging runs; however, this benefit appeared to be balanced by the acquisition of more images during continuous imaging runs.

Overall, these results suggest that with regard to sensitivity, both sparse and continuous protocols perform adequately. However, despite the contradictory outcomes, there is evidence that the acoustic spectrum of the scanner noise influences the relative abundance of characteristic frequencies. This will be a minor issue when qualitative maps are pursued (e.g. to outline cortical fields) or when making comparisons of data that were obtained using the same protocol (e.g. to assess tonotopic reorganisation), in which case continuous scanning is defensible. However, in the absence of compelling evidence to the contrary, it should be regarded as a significant confound when making quantitative comparisons between sound frequencies (e.g. to detect over/underrepresented sound frequencies), in which case sparse scanning must be preferred.

3.3. Field strength and voxel size

The increased signal-to-noise ratio and available BOLD signal associated with ultra-high magnetic field systems (≥3 T) allows the use of smaller voxel sizes; and the BOLD signal is better restricted to cortical grey matter because the signal strength of blood in draining veins is reduced due to shortened relaxation time at higher fields (van der Zwaag et al., 2009, 2011). Higher-field systems also come with additional challenges including increased geometric distortions and increased physiological noise (van der Zwaag et al., 2009). The development of multi-channel array head coils for fMRI offers improved signal sensitivity at all field strengths. Fine-scale individual subject mappings of tonotopy (<4 mm³ voxel volumes) have been obtained with 7-T systems in the cortex (Formisano et al., 2003; Da Costa et al., 2011, 2013) and in the inferior colliculus (De Martino et al., 2013).

Fig. 3 assesses the effect of voxel resolution on mapping outcome. Shown are three spatial samplings of the same unsmoothed single-subject tonotopy data set acquired at 1.5 mm isotropic resolution at 7 T. The data is re-sampled to 1 mm, 2 mm, and 3 mm isotropic corresponding to voxel volumes of 1 mm³, 8 mm³, and 27 mm³, respectively. Resampling and data analyses were performed in volumetric space and mapping results were projected onto cortical surfaces with no spatial smoothing applied. As can be seen, some map details are lost at the 3-mm isotropic voxel size but the overall tonotopic pattern is maintained. This demonstrates that tonotopic maps should be attainable at near-standard fMRI voxel sizes, provided that reliable signal is provided by the scanning system. One caveat is that at lower spatial resolutions, the two medial high-frequency end points could appear to merge, particularly in those individuals for whom they are close together in volumetric space.

3.4. Discrete frequency conditions

Many tonotopic mapping studies present pure tones of different frequencies as separate conditions and evaluate responses to each condition using a general linear model (GLM) (Formisano et al., 2003; Humphries et al., 2010; Woods et al., 2009; Langers and van Dijk, 2012; Herdener et al., 2013). Typically, each frequency...
condition consists of a train of tone bursts of a specified frequency or tone bursts within a limited band of frequencies around a specified centre frequency. Each condition is represented as a separate regressor and the weights associated with each regressor are solved for based on a least squares fit to the data, on a per voxel basis. The weights (beta values) corresponding to different frequency conditions may then be compared to determine a voxel’s stimulus selectivity, with widely-accepted statistical methods developed for both single and group-level comparisons. To be solvable under the GLM framework, conditions are presented in pseudo-randomized order to avoid correlation between regressors. Recently, Langers and van Dijk (2012) applied novel data-driven analyses to the evaluation of GLM regressor weights and obtained robust tonotopic maps at individual and group levels.

3.5. The travelling-wave design

Phase-encoded or travelling wave methods have been shown to be highly efficient in visual retinotopic mapping (Engel, 2012; Wandell and Winawer, 2011), somatotopic mapping (Zeharia et al., 2012), and currently tonotopic mapping (Talavage et al., 2004; Striem-Amit et al., 2011; Da Costa et al., 2011; Barton et al., 2012). In this paradigm the mapped parameter is gradually cycled through a continuous range of values thus creating a travelling wave of neural activity across the surface of the topographic map. A voxel’s stimulus selectivity is revealed, not by the response amplitude, but by the response phase which may be estimated by Fourier analysis or cross-correlation analyses of individual voxel time-courses. Travelling-wave tonotopy studies have presented tones either in continuous sweeps (Talavage et al., 2004; Striem-Amit et al., 2011) or in fixed steps (Da Costa et al., 2011) through logarithmic frequency progressions. Each stimulus cycle should be sufficiently long to allow responses to return to baseline between successive response peaks. Improved results may be obtained by combining runs of forward and reverse order progressions to avoid biases related to stimulus order. For example, the beginning of each stimulus cycle is associated with an abrupt onset that will activate sound-responsive but weakly-selective voxels, irrespective of frequency. A silent gap may be introduced at the discontinuity between cycles that, if sufficiently long, would allow baseline response measurement.

Da Costa et al. (2011, 2013) obtained tonotopic mappings of A1 and R that were clear at the individual subject level requiring only 16 min of fMRI scan time, demonstrating the efficiency of the travelling wave technique. Dick et al. (2012) coupled the travelling wave design with a highly novel stimulus (bandpass swept non-linguistic vocalisations) that was designed to be more complex and attentionally engaging than standard tone stimuli. Throughout visual cortex, retinotopic responses are enhanced by attentive view and, at higher-level areas, retinotopic responses depend strongly upon attentive viewing (Saygin and Sereno, 2008; Bressler and Silver, 2010). In the auditory cortex, attention appears to have widespread effects on primary as well as non-primary areas, including frequency-specific effects that colocalise with tonotopic frequency representations (Paltoglou et al., 2009; Da Costa et al., 2013).

3.6. Modelling of tuning widths

Going beyond maps of preferred frequency, recent studies have aimed to model the full response profile, or tuning width, of auditory cortex voxels. The studies of Moerel et al. (2012) and De Martino et al. (2013) were additionally novel in the use of natural sound stimuli (vocal, environmental, and tool sounds) for tonotopic mapping. Natural stimuli were represented by their spectral profiles (across 40 frequency bins) and regularized regression was used...
to estimate each voxel’s response to all spectral bins. Gaussian fits to the resulting response profiles allowed estimation of voxel–wise preferred frequency and tuning width. Maps of frequency preference obtained with natural stimuli were very similar to those obtained from separate scan runs with more standard pure tone mapping stimuli. Building upon a ‘population receptive field’ modelling approach from visual retinotopic mapping (Smith et al., 2001; Dumoulin and Wandell, 2008), Thomas et al. (2012) modelled the selectivity of each voxel as 1-dimensional Gaussian with a centre corresponding to its best frequency and a standard deviation indicating the tuning width. Best-fitting parameters for each voxel were obtained by fitting measured time-courses to a model time-course obtained by multiplying the stimulus sequence with the Gaussian population receptive field and then convolving with a measured haemodynamic response function.

These modelling techniques are advantageous in allowing a wide range of stimulus types without imposing constraints on stimulus order, and provide potentially useful estimates of tuning width. Single-neuron recordings in the monkey indicate that core neurons are more narrowly tuned than surrounding belt neurons (Rauschecker et al., 1995) leading to the idea that voxel tuning widths could help identify human auditory core. However, it is not straightforward to extrapolate single neuron tuning width data to human fMRI data, since each voxel represents the BOLD population response of hundreds of thousands of neurons collectively across cortical layers and neuronal types (10^5–10^6 neurons per cubic mm in cortex). Narrow frequency tuning in A1 is found primarily in the principal neurons of middle cortical layers and is degraded in superficial and middle cortical layers (Guo et al., 2012). At the population level, a heterogeneous mixture of neurons with varying frequency-tuning profiles would appear to have a broader bandwidth than individual neurons. The local field potential is a population-level electrophysiological signal that correlates better with BOLD responses than single-neuron recordings (Logothetis et al., 2001). A recent study in monkey auditory cortex, measuring cortical field potentials with microECOG, did not find a clear distinction in the sharpness or strength of frequency tuning between belt and core (Fukushima et al., 2012). To date, the proposed auditory tuning width maps from human fMRI data are complex and intriguing, and it remains to be seen if repeated studies provide convergent views.

3.7. Parcelling auditory cortex

Since the early days of fMRI retinotopic mapping emerged a widely-accepted tool and became a standard parcellation step for studies investigating visual cortical function. Tonotopic mapping can provide similar benefits to the study of auditory cortical function. Da Costa et al. (2013) recently used tonotopic mapping with high-field fMRI as an initial parcellation step to identify primary fields A1 and R at the individual subject level, and then showed that those primary fields were modulated by a frequency-specific auditory attention. Oh et al. (2013) identified A1 and R and found that those fields were modulated by auditory imagery. In summary, tonotopic mapping can be a useful first step toward further study of the function of human auditory cortex.

4. Conclusion

In conclusion, there is considerable consensus on the spatial distribution of frequency tuning in the region of auditory cortex surrounding HC in humans. Our interpretation of the existing literature supports a spatial orientation of PAC in humans that is consistent with that of non-human primates. Nevertheless, there is room for diverging interpretations in terms of cortical fields, especially in combination with other recently emerged neuro-imaging outcomes that are thought to be indicative of PAC. The reconciliation of these findings will remain an interesting area of investigation. At the same time, the latest developments in fMRI acquisition and analysis methods allow an ever more detailed investigation of tonotopy in the entire classical auditory pathway in humans, from subcortical auditory nuclei to non-primary auditory cortex, both in health and disease.

References


