CellPress

Feature Review
Roles of Supplementary Motor Areas in Auditory Processing and Auditory Imagery

César F. Lima,[1] Saloni Krishnan,[2] and Sophie K. Scott[1],*

Although the supplementary and pre-supplementary motor areas have been intensely investigated in relation to their motor functions, they are also consistently reported in studies of auditory processing and auditory imagery. This involvement is commonly overlooked, in contrast to lateral premotor and inferior prefrontal areas. We argue here for the engagement of supplementary motor areas across a variety of sound categories, including speech, vocalizations, and music, and we discuss how our understanding of auditory processes in these regions relate to findings and hypotheses from the motor literature. We suggest that supplementary and pre-supplementary motor areas play a role in facilitating spontaneous motor responses to sound, and in supporting a flexible engagement of sensorimotor processes to enable imagery and to guide auditory perception.

From Action to Sound

The premotor areas of the medial frontal cortex form a central node of the action network. More than a century ago, Horsley and Schafer [1] showed that electrical stimulation of these areas in monkeys produces movements of the trunk, proximal upper extremity, and head. The term 'supplementary motor area' was introduced in the 1950s to refer to this portion of the cortex [2], and although it was historically defined as a single area, more recent anatomical accounts suggest that it comprises two distinct fields: one anterior, the pre-supplementary motor area (pre-SMA), and one posterior, the supplementary motor area proper (SMA), with striking similarities in humans and monkeys [3–5] (Figure 1). These areas are known to be crucial for multiple aspects of motor behavior, including action preparation, initiation and selection of actions, motor learning, inhibition, conditional action, action control, and monitoring of action outcomes ([6–8], for reviews [9–11]).

In this review we address the candidate roles that SMA and pre-SMA play in auditory processing. These regions are commonly activated in auditory perceptual and auditory imagery studies, across a wide range of sounds including speech, nonverbal vocalizations, and music [12–17]. Notably, these responses are found even when the tasks do not involve overt motor components [14–17] (Figure 2). The role of this activity remains an unresolved issue, possibly because SMA and pre-SMA are traditionally conceptualized as being linked to action-related processes, unrelated to audition. Likewise, in auditory cognitive neuroscience, these regions do not form part of typical auditory processing networks, in contrast to temporal, inferior parietal, lateral premotor, and prefrontal areas (Box 1).

Anatomy and Motor Functions

SMA and pre-SMA are located anterior to the leg representation of the primary motor cortex, and they form part of the frontal agranular cortex, characterized by the lack of a granular layer IV [4,11] (Figure 1). Falling within the superior frontal gyri, they roughly correspond to medial

Trends

Hearing and imagining sounds–including speech, vocalizations, and music–can recruit SMA and pre-SMA, which are normally discussed in relation to their motor functions.

Emerging research indicates that individual differences in the structure and function of SMA and pre-SMA can predict performance in auditory perception and auditory imagery tasks.
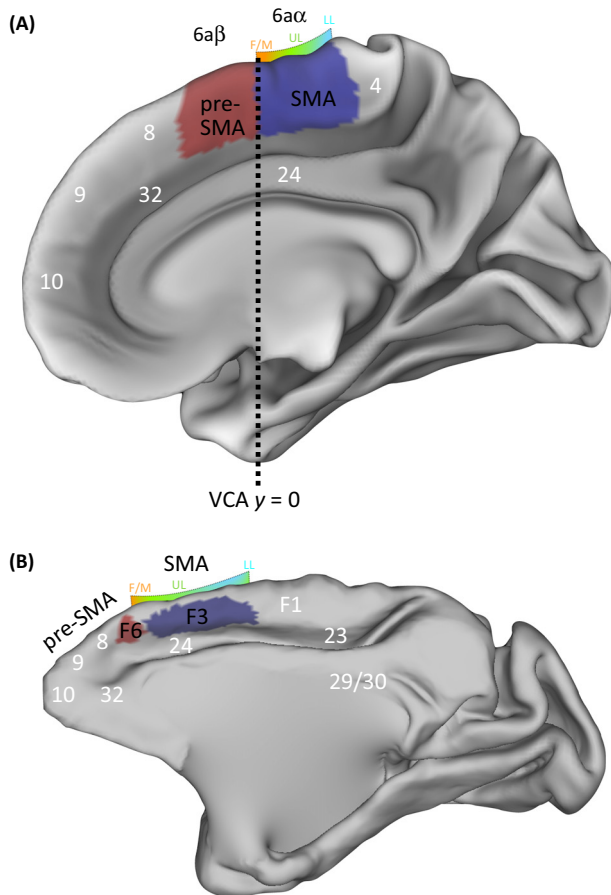
Responses during auditory processing primarily peak in pre-SMA and in the boundary area between pre-SMA and SMA. This boundary area is crucially involved in the control of speech and vocal production, suggesting that sounds engage this region in an effector-specific manner.

Activating sound-related motor representations in SMA and pre-SMA might facilitate behavioral responses to sounds. This might also support a flexible generation of sensory predictions based on previous experience to enable imagery and guide perception.

[1]Institute of Cognitive Neuroscience, University College London, London, UK
[2]Department of Experimental Psychology, University of Oxford, Oxford, UK

*Correspondence:
sophie.scott@ucl.ac.uk (S.K. Scott).

Brodmann area 6. Based on cytoarchitectonic and neurochemical differences, medial area 6 was later shown to comprise two distinct areas: SMA proper is located in medial area 6aα of Vogt and Vogt [18], and pre-SMA is located anteriorly in medial area 6aβ. Compared to SMA, pre-SMA shows a more-pronounced lamination and a clear demarcation of layer III from V [19], the cell density in the superficial layers is more homogeneous, and layer V is prominent and homogeneous [20]. In SMA, by contrast, layer III fuses with a dense layer Va [20]. The neurochemical footprint of these regions closely approximates these cytoarchitectonic differences, with clear stepwise changes in receptor binding densities at the boundaries between regions, and little variation within them [21]. Cytoarchitectonic and neurochemical similarities suggest that the SMA and pre-SMA in humans are homologous to macaque areas F3 and F6, respectively [10,19].

The border between SMA and pre-SMA in humans is considered to be the vertical plane through the anterior commissure (VCA line, $y = 0$) [11,22], and their inferior limit is at the level of the cingulate sulcus [23,24]. The anterior boundary of pre-SMA corresponds to the anterior end of the medial agranular cortex, bordering dysgranular cortex [4,25]. In neuroimaging work, because there are no macroanatomical landmarks for the anterior boundary, pre-SMA has been considered to extend to a virtual line passing through the genu of the corpus callosum [11].

Key Figure

Responses in SMA and Pre-SMA During Auditory Processing and Auditory Imagery
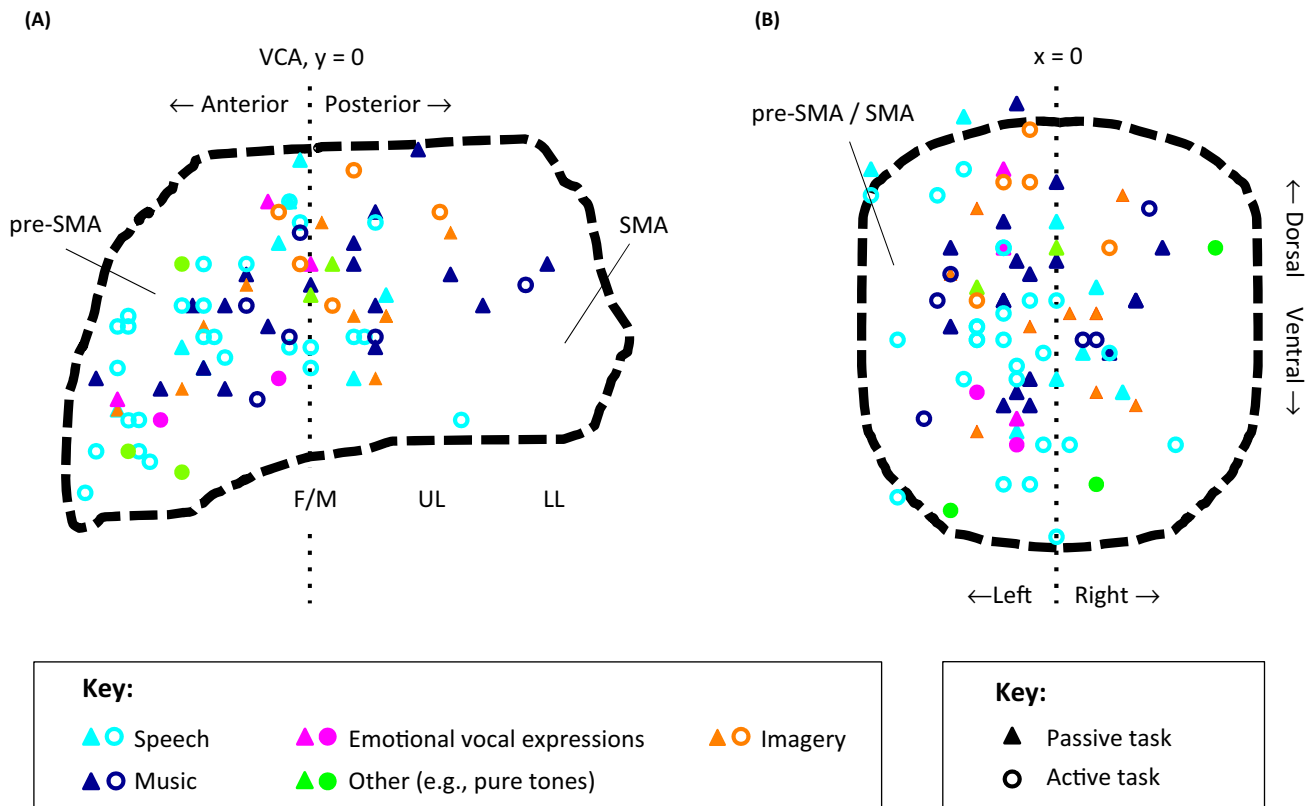
Figure 2. Sagittal (A) and coronal (B) views of peak auditory responses in SMA and pre-SMA, according to sound category and task type. The peaks (*n* = 83) were selected from 57 neuroimaging studies of auditory processing and auditory imagery (see online supplemental information for a list of the included studies). We have included studies using fMRI, PET, and voxel-based morphometry. Triangles denote activations obtained from tasks not including motor responses in contrast-relevant trials. Circles denote activations obtained from tasks including motor responses, such as button presses. We have included peaks derived from whole-brain level analysis that resulted from clear within-subject functional comparisons in healthy adult participants; that reflected categorical comparisons or a parametric effect of auditory-relevant variables; and, in the case of studies using an active task, inclusion was restricted to peaks resulting from comparisons between conditions involving similar motor demands. The extent and shape of SMA and pre-SMA were determined using the Human Motor Area Template (HMAT), which was created combining a meta-analysis of 126 functional studies with anatomical guidelines [26]. Peak voxel locations are plotted in Montreal Neurological Institute (MNI) space. Together, these data suggest that peak auditory responses tend to be seen in pre-SMA, and more often they cluster around the boundary area between pre-SMA and SMA. Only rarely activations extend posteriorly to SMA proper (A). In terms of lateralization, peak activations can be found in both hemispheres, but there is a predominance of findings in the left hemisphere (B). There is no evidence for spatially distinct patterns of activation across sound categories, or between auditory perception and auditory imagery. In (A) VCA refers to the vertical plane through the anterior commissure; and the location of the different body parts (F/M, face/mouth; UL, upper limb; LL, lower limb) illustrates the SMA somatotopic map in this region along an anterior-to-posterior orientation (imagery was not divided into different types of stimuli, i.e., speech and music, because of the relatively low number of studies for each type).
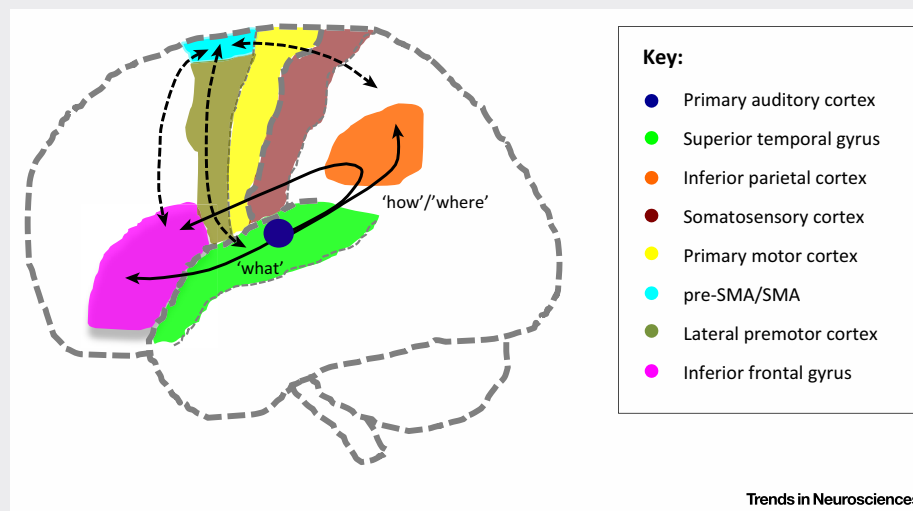
An activation likelihood estimation (ALE) meta-analysis indicates that the functional limit is located 12 mm anterior to the VCA line at *z* = 70, and 26 mm anterior to the VCA line at *z* = 40 [26]. SMA and pre-SMA are anatomically and functionally distinct from other motor and premotor areas, namely from the dorsal and ventral premotor cortices in lateral Brodmann area 6 [10,27–30].

Box 1. Streams of Auditory Processing in the Human Brain

As in the visual system, it is now well established that there are different anatomical and functional streams of processing in the auditory system. According to the dual-stream framework, sounds are processed in parallel along two largely segregated streams, an anteroventral and a posterodorsal one, both projecting from the primary auditory cortex (Figure I). These streams of processing engage with auditory information in distinctly different ways, and they play complementary roles in perception. The anteroventral pathway is important for the recognition and identification of auditory objects ('what'), and it forms a hierarchically organized stream, involving parts of the auditory association cortex in the superior temporal gyrus and the inferior frontal gyrus. The posterodorsal pathway is important for aspects of sensorimotor integration and spatial processing ('how/where'), involving posterior superior temporal fields, the inferior parietal cortex, motor and sensory areas, and the inferior frontal gyrus (for reviews [119–121]).

The majority of studies on the perception of linguistic information in sound (e.g., intelligibility in speech as compared to non-intelligible sounds) show recruitment of anterior auditory 'what' fields, consistent with a move towards more abstract semantic representations in the temporal poles [122,123]. By contrast, posterior auditory 'how' areas seem to code both auditory and sensorimotor information. They are recruited not only during listening to sounds but also during speech production and silent articulation [124,125], and in situations where articulatory behavior is modified via delayed auditory feedback [99]. In passive perceptual tasks, when heard sounds are meaningless (e.g., non-words, non-speech mouth sounds), posterior fields show greater activation than anterior ones [126]. There is also evidence that the posterodorsal pathway is strongly involved in the automatic, accurate control of auditory-motor actions. For example, listeners are able to automatically track subtle timing changes in sound sequences when they are asked to tap along to the sequences, and therefore engage posterodorsal systems; if they are asked about these changes, however, they fail to report hearing them, suggesting that the sensorimotor adjustments were controlled below the level of awareness [127].

Auditory processing is therefore a multifaceted phenomenon, involving multiple mechanisms that can interact in a complex manner. Understanding the role of SMA and pre-SMA in relation to these mechanisms might be especially relevant because action and sound are intimately related, and these regions are known to be connected to central nodes of the auditory network, namely inferior frontal, temporal, and parietal regions (Figure I) [20,40,44–49].

Key:
- ● Primary auditory cortex
- ● Superior temporal gyrus
- ● Inferior parietal cortex
- ● Somatosensory cortex
- ● Primary motor cortex
- ● pre-SMA/SMA
- ● Lateral premotor cortex
- ● Inferior frontal gyrus

Trends in Neurosciences

Figure I. Streams of Auditory Processing in the Human Brain. Lateral view of the anteroventral and posterodorsal auditory processing pathways (unbroken arrows), and anatomical connections between SMA and pre-SMA and inferior frontal, temporal and parietal areas (broken arrows). The main anatomical systems involved in sound processing are shown in color. Projecting from the primary auditory cortex into anterior temporal and inferior frontal areas, the anteroventral stream decodes meaning in sounds ('what'). The posterodorsal pathway supports sensorimotor and spatial processes ('how' and 'where'), and it projects posteriorly into inferior parietal, sensory, motor, and inferior frontal areas.

SMA and pre-SMA differ in terms of their motor functions. SMA is tightly coupled with movement generation and control. Stimulation, neuroimaging, and lesion work in humans and non-human primates reveals a somatotopically organized map in this region along an anterior-to-posterior, face-to-legs gradient. Orofacial movements, including speech production, are associated with

sites located anteriorly, at the border with pre-SMA, whereas upper and lower limb movements are associated with more posterior sites [1,31–36] (Figure 1). Similar somatotopic maps are found in motor and lateral premotor cortices, running in a ventral-to-dorsal, face-to-legs gradient [28,30,36,37]. In contrast to SMA, motor responses in pre-SMA can only be evoked using high current intensities, and the association between stimulation and movements is more variable. In humans and non-human primates, such evoked movements are typically slow, complex, and mainly involve orofacial and upper limb effectors [5,11,32,34]. Pre-SMA has been suggested to support higher-order aspects of action, including action preparation, the organization of movement sequences, the perception of objects that are potential targets of motor acts, and responses to sensory stimulation during cued movements [3,5,10,38]. The distinct functional properties of SMA and pre-SMA are potentially related to their patterns of connectivity.

Connectivity analyses show that SMA, but not pre-SMA, can directly influence motor output via its direct connections to the primary motor cortex and spinal cord [20,22,39,40]. This region is also connected with the posterior premotor and cingulate areas, and parietal regions [20]. By contrast, pre-SMA is richly connected to prefrontal, anterior premotor, and cingulate areas, as well as to the medial parietal cortex [20,22,41]. Given its position between the prefrontal and motor areas, pre-SMA can be conceptualized as an area that integrates information about action plans, motivation, and objects, which can be used for action initiation and control [10]. Both SMA and pre-SMA receive basal ganglia projections [40,42,43], with pre-SMA connecting to more anterior sites of the striatum than SMA, as indicated by *in vivo* imaging tractography in humans [22,43]. Such corticostriatal connections have been suggested to be part of a wider network that, through the thalamus, projects back to the cortex, and supports different aspects of motor control [40]. Furthermore, evidence from tracer injection in non-human primates also indicates that pre-SMA is connected to SMA [20].

Importantly, there are structural and functional connections between SMA, pre-SMA, and regions that are considered to be central for sound and speech processing (see Figure I in Box 1). Post-mortem dissection and imaging tractography evidence from humans has revealed a direct pathway, the frontal aslant tract, directly connecting SMA and pre-SMA with the opercular part of inferior frontal gyrus [40,44], which corresponds to Broca's area in the dominant hemisphere and is a core component of the posterodorsal auditory pathway. The frontal aslant tract is observed in both hemispheres, but diffusion tractography indicates that it is left-lateralized in most right-handed individuals [44]. A study of primary progressive aphasia found that reductions in the structural integrity of this tract relate to performance in a verbal fluency task [45], suggesting an involvement in language production. Pre-SMA also has connections to the temporal lobe. Tracer injection studies in non-human primates have revealed connections to the anterior superior temporal sulcus [20,46], and functional connectivity analyses in humans reveal connections to the superior temporal gyrus [47,48]. Both SMA and pre-SMA are connected with the parietal lobe, including inferior parietal sites [20,48,49], which are part of the posterodorsal pathway (see Figure I in Box 1).

Auditory Information in SMA and pre-SMA

Consistent with their patterns of connectivity, SMA and pre-SMA are recruited during auditory processing (Figure 2, Key Figure). For speech perception, functional neuroimaging studies have reported activity in response to syllables [50,51], words [52,53], and sentences [54–59]. Such responses seem to be modulated by the difficulty of speech processing and comprehension. In pre-SMA, activity is inversely correlated with speech intelligibility as manipulated by background noise; in other words, responses are stronger when the speech signal is less intelligible and comprehension is more challenging [15,60]. In an ALE meta-analysis of functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) studies, Adank [61] observed that, in addition to anterior insula fields, pre-SMA was the only area of the motor

---

Box 2. Emotional Voices in SMA and Pre-SMA

The processing of nonverbal vocal emotional cues is key for vocal communication. These cues include purely nonverbal vocalizations, such as laughter or crying, and modulations of the tone of voice while speaking (emotional prosody). Several studies indicate that listening to vocal emotional cues, and inferring their meaning, recruit SMA and pre-SMA. In a study of nonverbal vocalizations, including laughter, triumph sounds, screams, and disgust sounds, Warren *et al.* [14] showed that responses in SMA and pre-SMA are significantly modulated by the type of sound, in a passive listening task not involving overt motor responses. These modulations were further related to broader affective dimensions of arousal, with high-arousal vocalizations activating these regions more strongly than low-arousal ones. While these findings were taken to reflect a spontaneous impulse to respond to the emotional expressions of others, more recent work showed that responses in these regions might also contribute to perceptual processes. McGettigan *et al.* [16] reported activations in pre-SMA during passive listening to voluntary social-type laughter and spontaneous laughter, and found that the magnitude of these responses predicted performance in a post-scanner behavioral authenticity detection task. Consistent with a role in the perception of auditory expressions, Bestelmeyer *et al.* [12] showed that pre-SMA responses are associated with greater discrimination between expressions taken from continua of anger and fear nonverbal vocalizations, even after regressing out the low-level acoustic features of the expressions.

Activations in SMA and pre-SMA can also be obtained for the perception of emotional prosody signals during passive listening [140] and during explicit evaluations of emotional authenticity [141]. There is suggestive evidence that this might be modulated by training. Kreifelts *et al.* [142] showed that 4 weeks of non-verbal emotion communication training produces selective changes in a network of regions including bilateral pre-SMA. This was observed in a task requiring valence judgments of multimodal stimuli comprising speech prosody and facial emotional cues.

The links between SMA and pre-SMA and vocal emotional processing possibly reflect a broader association between these regions and socio-emotional variables. Resections of the pre-SMA correlate with impairments in the ability to infer intentions (mentalizing) [143], and a transcranial magnetic stimulation study showed that stimulation of SMA modulates the perceived valence of emotional visual stimuli [144]. The interactions between motor-related activity in SMA and pre-SMA and socio-emotional factors could be supported by direct and indirect anatomical connections between these regions and emotional systems including cingulate regions and the limbic system [40,144].

---

and premotor cortices recruited during the processing of distorted/less-intelligible speech as a result of background noise, regional accents, or speech rate differences.

The involvement of SMA and pre-SMA is not limited to the perception of speech: activity is observed during listening to familiar [13,17] and unfamiliar music [62,63], as well as during the perception of action sounds (e.g., drinking from a straw) [64], and sequences of tones [65,66]. Speech melody and nonverbal emotional vocalizations, such as laughter and crying, also evoke responses in SMA and pre-SMA [12,14,16,67], which suggests that auditory responses in these regions extend to the processing of socio-emotionally salient sounds (Box 2).

SMA and pre-SMA are consistently implicated in auditory imagery as well, when participants are instructed to generate auditory mental images in the absence of sensory input (Box 3). This includes imagery of speech, such as syllables and words [68,69], and imagery of music ([13,70]; for a meta-analysis [71]). Indeed, these regions arguably predominate over sensory cortex in such auditory imagery tasks: studies directly contrasting music listening and music imagery showed that SMA and pre-SMA, unlike motor and lateral premotor cortices, are recruited more strongly for imagery than for listening, while the superior temporal gyrus is recruited more strongly for listening than for imagery [13,72,73].

Not all forms of auditory imagery are voluntary: people often experience involuntary musical imagery, a phenomenon colloquially termed 'earworms' (Box 3); and auditory verbal hallucinations form another example of auditory images that are experienced as uncontrollable and as arising from an external source. Pre-SMA might be important for the sense of volitional control (intentionality) over auditory images. In an fMRI study comparing auditory verbal hallucinations and voluntary imagery of the same voices, pre-SMA activity was stronger during imagery than during hallucinations, both when hallucinations were compared with imagery within the same group of patients and when they were compared with imagery in healthy controls [74]. Work on

CellPress

Box 3. Auditory Imagery

Auditory imagery refers to the process by which individuals generate auditory information in the absence of sound perception, such as when we imagine the voice of a friend or the sound of a familiar song. Auditory imagery can be reported to be so vivid that it resembles the experience of hearing, and to be as accurate as auditory representations arising from sensory input [115]. Auditory mental images cannot be directly observed or measured, but their properties have been inferred using strategies such as self-report [13,80,114], and performance-based tasks in which participants provide judgments that require engaging in imagery (e.g., judging the pitch of words taken from familiar tunes in the absence of auditory input [73]; or judging whether a final note of a scale is mistuned when the initial notes were played but the remaining ones had to be imagined [115]). An approach often used in functional neuroimaging and electrophysiological research consists of comparing conditions where participants generate auditory imagery with conditions where they are presented with auditory stimuli, or engage in tasks not involving imagery [13,69,72,73,81,115].

Studies focusing on voluntary imagery of music and speech reveal the recruitment of a network of brain regions including the superior temporal gyri, parietal, motor, and lateral premotor cortices, the inferior frontal gyrus, and SMA and pre-SMA ([13,68–70,72,73,80,81]; for a meta-analysis [71]). This network substantially overlaps with the network recruited during auditory perception, including auditory areas. Auditory cortex activity during imagery is thought to reflect the instantiation of sound-like representations as controlled by higher-order cortical mechanisms [13,73], contributing to the phenomenological experience of 'hearing'.

Little is known about involuntary forms of auditory imagery, such as 'earworms' (the experience of having music looping into one's head). Farrugia et al. [145] reported associations between self-report frequency of earworms and cortical thickness in frontal and temporal cortices, and in the anterior cingulate and angular gyrus. The length of earworm episodes correlated with cortical thickness in anterior pre-SMA, and affective evaluations of this form of imagery correlated with grey matter volume in temporal-polar and parahippocampal cortices.

Studies on mental imagery are typically confined to a single modality (e.g., auditory, visual, motor), but there is some evidence suggesting that, while superior temporal areas play an auditory-specific role, other components of the auditory imagery network, namely motor and premotor systems, might be also engaged by other forms of imagery [71,80]. More research directly contrasting different forms of imagery will be needed to address this question.

Box 4. Individual Differences and Plasticity in SMA and Pre-SMA

The structure and function of SMA and pre-SMA can vary between individuals as a consequence of motor learning and expertise. Structural changes can occur over relatively short timescales–for instance, grey-matter increases in these regions were observed over 6 weeks of practicing a complex motor skill (dynamic balance task), and these structural changes correlated with changes in functional connectivity over time [128,129]. Structural changes in SMA have also been found in specialized groups of participants who completed intense long-term sensorimotor training, such as ballet dancers [130] and athletes [131,132]. Directly relevant to auditory processing, a large study of musicians [133], who spend years developing a repertoire of specific motor actions in relation to their sensory consequences, revealed a positive correlation between musical expertise and increased grey matter volume in SMA.

In addition to structural changes, motor skill learning can modulate functional responses in SMA and pre-SMA–in the context of initial stages of learning, activity has been shown to decrease in pre-SMA and to increase in SMA as performance improves ([134] for review). Long-term auditory-motor experience also correlates with functional changes in SMA. For instance, opera singers show greater activity in left SMA during overt singing than novices do [135]. These changes might explain why musicians perform better in new motor tasks. While they might be able to automatically recruit SMA and pre-SMA for motor performance and sensorimotor integration, independently of the complexity of the task, the engagement of non-musicians might be more strongly constrained by task difficulty [136,137].

Not only does functional activity in SMA and pre-SMA appear relevant to action execution but it may also shape the auditory perception of learnt actions. When listening to music, musicians recruit pre-SMA more strongly than non-experts, in a way that does not seem to reflect a general unspecific effect–it is selective for music that they can produce [79]. Furthermore, higher sensitivity to beat perception in music relates to a stronger engagement of SMA [138]. Interestingly, SMA activity during auditory processing can be modulated by other forms of sensorimotor expertise in addition to musical expertise. Athletes who play basketball or tennis show stronger activity in SMA during listening to sport sounds, such as a racquet hitting a ball, as compared to non-sport sounds, such as paper crumpling. In addition, even within the category of sport sounds, SMA activations are stronger for sounds related to the sport that the athletes can play than to sounds related to a sport that they cannot play [139].

auditory hallucinations in healthy controls ('nonclinical voice hearers') has shown that pre-SMA is recruited both during voluntary imagery and hallucinations; notably, however, while during imagery pre-SMA becomes activated before inferior frontal and temporal fields, during hallucinations it becomes activated at the same time as these areas [75], suggesting an impaired control of auditory and motor processes by pre-SMA in hallucinations.

Although SMA and pre-SMA activity is obtained during auditory perceptual and imagery tasks, it could simply reflect task-related motor or general cognitive processes, such as button presses or response selection, or it could reflect an epiphenomenal response, unrelated to audition. However, there is strong evidence for such activity during passive listening to different types of sounds, when no movements or response selection are involved [14–17,54,76,77] (Figure 2). Indeed, even when movements are involved, such as during overt task performance, activations in these regions are can be seen over and above the neural responses associated solely with motor performance [12,67,78]. SMA and pre-SMA activity further varies with auditory-motor experience; studies on expert musicians indicate that auditory-motor training modulates responses to sound in these regions, pointing to a specific association beyond task effects [79] (Box 4). Consistent with a direct role for SMA and pre-SMA in auditory processes, inter-individual differences in the structure and function of these regions can predict outcomes in auditory imagery and auditory perception tasks. In a voxel-based morphometry study, larger grey matter volume in SMA was shown to predict higher perceived vividness of mental images during auditory imagery [80]. In a related vein, evidence from fMRI indicates that larger functional responses in SMA during imagery of familiar melodies correlates with higher perceived vividness of auditory images [81]. In auditory perception, the magnitude of functional responses in pre-SMA during passive listening to posed and spontaneous laughs was shown to correlate with better ability to categorize laughter authenticity in an off-line behavioral task [16]. In a task requiring participants to judge the direction of temporal changes in frequency-modulated sounds ('speeding-up' or 'slowing-down'), while ignoring simultaneous pitch changes in the same sounds, pre-SMA responses were linked with perceptual sensitivity to discriminate temporal changes, as well as with reduced susceptibility to distortions in temporal judgments induced by pitch changes [82]–in other words with a better ability to focus on the task-relevant features of the stimuli (temporal changes) and ignore the irrelevant ones (pitch changes). These findings provide initial support for the hypothesis that SMA and pre-SMA directly influence a variety of auditory processes. However, determining the nature of this contribution, and whether it represents a modulatory or a central contribution, requires further investigation. Studies of patients with SMA and pre-SMA lesions, and using techniques such as transcranial magnetic stimulation (TMS), will be important for this. TMS studies have shown that disrupting activity in primary motor and lateral premotor fields impairs the discrimination of speech sounds [83] and vocalizations [84]. This approach could be extended to examine the roles of SMA and pre-SMA in the perception and imagery of different types of auditory information (see Outstanding Questions).

In terms of lateralization of activity, many studies report bilateral involvement of SMA and pre-SMA during auditory processing [52,64,80,85], and peak activations can be found in both hemispheres. However, left hemisphere responses predominate, if we take the relative number of peak activations reported in the left and the right hemisphere as an indication of lateralization (Figure 2). This may mean that there is more to learn about hemispheric asymmetries within these regions (see Outstanding Questions). Given that assigning midline responses to the left or right hemisphere might be difficult using fMRI, future research will benefit from also using other approaches to address this question, such as single-cell recordings or studies of patients with unilateral lesions.

Figure 2 shows that peak auditory and imagery responses in SMA and pre-SMA are not evenly distributed across the surface of these regions. They are primarily found in pre-SMA, and in a

cluster around the border between pre-SMA and SMA; responses are more rarely found in posterior SMA. Interestingly, the area around the border between the two regions is consistently associated with the production of orofacial movements, vocalizations, and speech [31–33,35]. Lesions in this area can selectively produce aphasia, without motor limb impairments [31], and affect the production of vocalizations in squirrel monkeys [86]. Thus, in addition to engaging pre-SMA, auditory information engages SMA fields that are also involved in the generation and control of movements related to vocal production. This is suggestive of effector-specific motor recruitment during auditory processing and auditory imagery. Evidence for the spatial overlap between perception and production in these regions comes from several studies directly comparing the two conditions [14,16,56,64] (Box 5). It is also interesting to note that the frontal aslant tract, that connects the inferior frontal gyrus with medial frontal areas and has been suggested to play a role in language production [45], terminates at the border area between SMA and pre-SMA [40]. The possible involvement of this tract in auditory perception and imagery has yet to be determined.

Candidate Functions of SMA and Pre-SMA Responses

There is no consensus position on the roles of SMA and pre-SMA responses in auditory processing and imagery. When such responses are discussed, they have been linked to a variety of processes. Timing functions have been suggested for perceptual tasks requiring evaluations of temporal aspects of auditory stimuli [82], or for stimuli varying in the sequential predictability and rhythmic regularity that they afford [66,76,87–89]. SMA and pre-SMA, together with the cerebellum and the basal ganglia, have in fact been considered to form the substrates for a 'temporal processing' network (for reviews [90–93]). By contrast, sub-vocalization and

---

Box 5. Linking Auditory Perception and Production in SMA and Pre-SMA

There is strong evidence that SMA and pre-SMA have sensorimotor properties, and respond to both the perception and execution of actions. Neuroimaging studies report overlapping activity in the right and left pre-SMA during action perception and action production, as revealed by an activation likelihood estimation (ALE) meta-analysis [146]. Importantly, this overlapping activity does not simply reflect the averaging over voxels that respond to perception and production at a group level–instead, shared responses were shown to consistently emerge at a single-subject level [147]. Further evidence comes from single-cell recordings. Mukamel *et al.* [148] recorded extracellular activity in human medial frontal and temporal cortices, and found that a significant number of neurons in SMA responded during the execution and observation of hand grasping actions and facial emotional expressions.

Many studies focus on perception–production couplings in the visual domain, but there is evidence for similar couplings in the auditory domain. Gazzola *et al.* [64] observed overlapping SMA activity during the execution of motor actions and during passive listening to sounds associated with those actions. Overlapping activity has also been found in SMA and pre-SMA during listening to emotional vocalizations such as laughter, and during the production of related orofacial movements [14,16]. Studies on expert musicians suggest that these sensorimotor associations might be modulated by learning and experience–while both listening to and producing music lead to overlapping activity in SMA, this region is more strongly engaged by musicians than by non-musicians [149].

An emerging literature suggests that sensorimotor links in SMA and pre-SMA extend to speech processing. In an ALE meta-analysis of studies examining the neural bases of speech production and speech perception in difficult listening conditions, Adank [61] reported that SMA and pre-SMA are part of both speech comprehension and production networks. These regions are additionally involved in monitoring auditory feedback during speech production, particularly in demanding or noisy speaking conditions [100,150]. Gauvin *et al.* [56] observed that monitoring speech errors in both speech perception and production tasks leads to activity in the left pre-SMA. This suggests that SMA and pre-SMA are sensorimotor resources that can be recruited when speech processing becomes challenging, perhaps because they can be used to generate and constrain predictions about the sensory input.

The overlap of perceptual and action mechanisms in SMA and pre-SMA suggests that these regions are part of the 'mirror neuron' system. Research on mirror neurons has traditionally focused on ventral premotor and inferior parietal areas, but some studies have begun to discuss SMA and pre-SMA in the context of this system [64,146–148]. Important claims of the mirror neuron theory remain controversial, however, including the ideas that these neurons code action goals and directly support the understanding of actions and intentions [151].

articulatory/motor roles for SMA and pre-SMA have also been suggested, in the context of studies on imagery and perception of speech and music, because these stimuli can be easily reproduced using the vocal apparatus [13,15,69,73]. When tasks emphasize detecting abnormalities in the signal, such as speech errors or noise-related interruptions in speech [56,94], SMA and pre-SMA have been linked to monitoring functions. There are few attempts to link these different proposals, and a unified framework that accounts for perceptual and imagery processes across different sound categories has not been proposed.

Our starting point in developing a model of SMA and pre-SMA roles in auditory processing is our finding that peak auditory and imagery responses cluster in pre-SMA, and extend to the boundary between pre-SMA and SMA (Figure 2). This might indicate the involvement of higher-order control of action-related processes, together with coupling to vocal production, because more complex and controlled aspects of motor behavior are linked with pre-SMA, while a more direct role in action execution is linked with SMA [9–11]. Several higher-order functions have been attributed to pre-SMA in humans and non-human primates, namely the integration of sensory and motor information (sensorimotor processing), planning of actions, focus on current goals while inhibiting interference from irrelevant cues, and intentional initiation of action [9,10,95,96]. In the context of speech production, for example, fMRI and PET studies have shown that complex processes such as word selection and propositional speech recruit anterior pre-SMA sites, whereas simple articulatory movements recruit more posterior sites [97,98]. In the context of temporal processing, tasks emphasizing perceptual aspects of the stimuli (e.g., making temporal judgments of auditory stimuli) tend to recruit pre-SMA, whereas tasks emphasizing motor responses (e.g., tapping along to a rhythm) recruit relatively more posterior sites [93].

We argue that pre-SMA and the boundary area between pre-SMA and SMA could provide a mechanism for linking auditory information with related motor programs, and that this sensorimotor engagement could be modulated by controlled processes as part of a wider network involving prefrontal, auditory, and other sensorimotor systems. We propose two parallel candidate roles for SMA and pre-SMA: motor facilitation, which might provide a mechanism for supporting spontaneous behavioral responses to sounds; and a more controlled and flexible process, in that motor programs would be retrieved to generate and exploit sensory expectations that would enable imagery and optimize perceptual processes. Aspects of auditory perception and auditory imagery may be characterized by such auditory–motor interactions.

Auditory–Motor Interactions

Sound and action are inherently linked. Our auditory experience can be conceptualized as a sensorimotor interplay: the sounds that we hear reflect actions, and the sounds that we make result from actions. We learn to associate actions, such as speaking or singing, with their sensory correlates through experience. This establishes bidirectional links between actions and the corresponding sounds. One illustration of these associations is the role of auditory information in the control of speech production. Manipulations of the online perception of self-produced speech, for instance via delayed auditory feedback, detrimentally affect the act of speaking, and modulate activity in sensory and motor networks [99,100]. Music training with auditory rhythmic cueing improves motor functions in Parkinson's patients (gait kinematics), and such auditory cueing-related benefits extend beyond gait to motor timing tasks and to auditory time perception (duration discrimination of pure tones and detection of misaligned beats in music) [101]. SMA and pre-SMA are well placed, anatomically and functionally, to support interactions between motor and sensory processes in both an overt and a covert manner. Pre-SMA is recruited during auditorily cued actions, and such recruitment is stronger when participants learn to associate auditory cues with specific actions than when they listen to the same cues and produce the same actions without a specific association between them [102], suggesting a role in

sensorimotor integration. SMA and pre-SMA are also important for motor imagery [103,104], possibly relying on the mechanisms that also support overt motor execution [105]. Activating motor programs in these regions can thus serve action execution, as well as cognitive processes that do not involve overt motor behavior but might benefit from motor-related processing. Furthermore, motor information can modulate sensory processes via efference copies of the motor programs, sent in parallel to sensory systems and to relevant effectors, and SMA has been pointed out as a likely source of this modifying activity [47].

Spontaneous Engagement and Motor Facilitation

We suggest that SMA and pre-SMA are part of the network that supports the spontaneous engagement of motor programs related to sounds. Some sounds, owing to their rhythmical patterns or their social and motivational salience, elicit motor responses, such as singing, tapping, dancing, or vocal alignment/contagion. Infants aged 5–24 months spontaneously coordinate body movement to music, with the degree of coordination correlating with positive affect [106]. Vocalizations such as laughter can be primed solely by listening to someone laugh, and are strongly primed by the presence of other people [107]. During spoken interactions, people spontaneously synchronize breathing, coordinate turn-taking, and imitate others' speech patterns, for example in terms of intonation, speech rate, and accent [108–110]. Such propensity to respond to rhythmic, social, and emotional auditory information might promote social convergence, learning, coordination, and affiliation.

Electrical stimulation of SMA and pre-SMA can elicit a subjective 'urge' to move [32], and neural oscillations and fMRI studies indicate that these areas are engaged during passive listening and motor synchronization to auditory rhythms [76,111]. Consistent with this, SMA and pre-SMA are also activated during passive listening to music, and responses are stronger for rhythmic sequences as compared to random sequences [87], and for familiar music as compared to random sequences [77]. In an fMRI study, Silva-Pereira *et al.* [17] showed that SMA is more strongly recruited by familiar than unfamiliar music, and by music that listeners like as compared to music that they dislike. This suggests that the readiness to respond with motor activity to music might be affected by how much we can predict upcoming events based on previous knowledge, and that motivational factors might play a modulatory role as well. SMA and pre-SMA are also engaged during listening to emotionally salient human vocalizations, including laughter [14] (Box 2), which are primed by both social and emotional factors.

Controlled and Flexible Processing

In parallel with a role in facilitating spontaneous behavioral responses, SMA and pre-SMA are well placed for supporting a more controlled engagement of sensorimotor processing during auditory perception and imagery. Pre-SMA, in particular, as an interface area between prefrontal and motor systems, could support the initiation and control of sensorimotor engagement, integrating information about goals, context, and motivation in that process. This would provide a mechanism for generating internal estimates of the sensory correlates of sound-related actions, and these estimates would enable dynamically updated sensory expectations that could flexibly contribute to various aspects of perception and imagery. In the context of imagery, estimates of the sensory correlates of candidate actions would be the substrate of the subjective experience of 'hearing'. This is consistent with the observation that responses in the superior temporal gyrus and in pre-SMA and SMA are both seen in auditory imagery studies [13,69,71], and with evidence linking individual differences in the structure and function of SMA with higher vividness of auditory images [80,81]. Further evidence for a role of pre-SMA in controlled/ intentional aspects of imagery comes from studies showing that this region is more strongly engaged, and is activated before auditory regions, during voluntary imagery as compared to (involuntary) auditory verbal hallucinations [74,75]. The involvement of SMA and pre-SMA in activating relevant motor programs (motor simulation) during auditory imagery, and the potential

contribution of this for estimating and predicting auditory representations, was recently suggested in the context of imagery of speech [69].

In auditory perception, the same controlled engagement of sensorimotor processes might contribute to guiding and optimizing different perceptual processes. Sensory expectations generated based on previous sensorimotor experience might orient attention to goal-relevant features of the auditory input, suppress irrelevant ones, reconstruct missing information, and monitor the flow of incoming events. These ideas resonate with the 'emulation' framework discussed in relation to action control, and generalized to motor imagery, perception, and prediction [112,113]. Converging mechanisms for auditory imagery and auditory perception in SMA and pre-SMA have been suggested [13,73,80]. For example, individuals reporting highly vivid auditory imagery show more specific neural representations of sounds during auditory perception in the same SMA peak that is associated with imagery [80]. Behavioral and electrophysiological studies further show that generating internal representations and expectations of auditory events benefits vocal production [114] and auditory perceptual processes [115,116]. In an fMRI study examining the perception of artificially degraded words, Shahin *et al.* [94] found stronger pre-SMA responses when participants subjectively experienced physically interrupted words as continuous, as compared to when they had the same experience of continuity for actual physically continuous words. By providing a flexible mechanism for engaging sensorimotor processes, this region might thus be involved in 'repairing' degraded auditory information. Such repair mechanism could relate to findings on other forms of degraded/distorted speech: pre-SMA is more strongly activated during the processing of time-compressed as compared to normal speed speech [85], and for speech in higher as compared to lower levels of background noise [15].

Auditory perception involves multiple streams of processing, and auditory objects can often be categorized via obligatory hierarchical processing in the anteroventral 'what' auditory pathway (Box 1). By contrast, the engagement of sensorimotor mechanisms in SMA and pre-SMA likely does not reflect early perceptual processing for meaning. Instead, it could be part of a parallel, more flexible stream of processing, and the magnitude of its contribution to perception and imagery could be a function of previous sensorimotor experience relevant to the listening/imagery context, difficulty of sensory-based processing, and task goals. A stronger contribution could be expected when previous experience enables the generation of rich and fine-grained sensory expectations, when the representation of the sensory input is degraded/distorted or absent (imagery), and when the task benefits from optimal sensorimotor integration or requires explicit/attentive access to auditory representations. Different lines of work provide initial support to this idea. Studies on expertise indicate that auditory-motor experience modulates SMA and pre-SMA responses during auditory processing [79] (Box 4), possibly reflecting a facilitated generation of sensory expectations. Stronger responses are obtained during auditory imagery, when there is no sensory input, as compared to listening [13,72,73]; and they are stronger during listening in difficult conditions, when sensory input is degraded or distorted, than when the sensory input is relatively intact [15,85,94]. Finally, there is a positive association between the structure and activation of SMA and pre-SMA and performance on a variety of explicit auditory tasks, requiring not only accessing temporal representations of the stimuli, but also judging other aspects such as socio-emotional categories and vividness of mental images [12,16,80–82].

An emerging theme from our review of the literature is that auditory information often engages SMA fields that are also linked to vocal production. Our vocal apparatus is arguably our primal, original instrument for making sounds. Because of the multifaceted and flexible nature of vocal production, the representation of these effectors could form a representational space within which we could generate sensorimotor estimations of different properties of sound. This could include the identity of linguistic information, and inferences about nonverbal signals such as

voices and music and how they unfold in time. It could also include aspects related to rhythm, timing, or sequencing of auditory events, because these are all central features of vocal production. For instance, Wolfensteller *et al.* [117] showed that SMA is recruited during rhythm predictions in visual sequences of abstract stimuli, as compared to predictions about positions and objects. This finding in the visual domain might reflect the correspondence between the emphasis of the task (rhythm) and a feature that is central for a given effector (articulatory system). Consistent with this, tracking temporal and rhythmic aspects of a variety of auditory and visual stimuli has been shown to recruit the boundary area between pre-SMA and SMA [93], similarly to vocal production.

Concluding Remarks

Auditory stimuli, both heard and imagined, activate SMA and pre-SMA. The broad evidence for these activations suggests that we in auditory cognitive neuroscience need to conceptualize these regions as part of the functional neuroanatomy of auditory processing, and to account for the variety of listening and imagery contexts in which they are involved. This matters because SMA and pre-SMA are not typically included in models of auditory processing and, when the roles of the motor and premotor systems are considered, the focus is primarily on lateral fields. We argue that SMA and pre-SMA support the activation of sound-related motor representations during auditory perception and auditory imagery, and that this activation can be modulated by controlled processes. According to our hypothesis, these regions mediate spontaneous motor responses to sound, and support a more controlled generation of sensory predictions based on previous sensorimotor experience, predictions that can be flexibly exploited to enable imagery and optimize a variety of perceptual processes. Future research will need to delineate the precise relationship between SMA and pre-SMA and the regions traditionally considered to the part of auditory streams of processing (Box 1). Another crucial aspect of future developments will be to refine our understanding of which aspects of auditory processing, context, and current goals/task are more directly dependent on the functions of SMA and pre-SMA.

Appendix A Supplemental Information

Supplemental information associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.tins.2016.06.003.

References

1. Horsley, V. and Schafer, E.A. (1888) A record of experiments upon the functions of the cerebral cortex. *Philos. T. Roy. Soc. B* 179, 1–45

2. Penfield, W. and Welch, K. (1951) The supplementary motor area of the cerebral cortex: a clinical and experimental study. AMA Arch. *Neurol. Psychiatry* 66, 289–317

3. Matsuzaka, Y. *et al.* (1992) A motor area rostral to the supplementary motor area (presupplementary motor area) in the monkey: neuronal activity during a learned motor task. *J. Neurophysiol.* 68, 653–662

4. Tanji, J. (1994) The supplementary motor area in the cerebral cortex. *Neurosci. Res.* 19, 251–268

5. Rizzolatti, G. *et al.* (1990) Neurons related to reaching-grasping arm movements in the rostral part of area 6 (area 6aβ). *Exp. Brain Res.* 82, 337–350

6. Bonini, F. *et al.* (2014) Action monitoring and medial frontal cortex: leading role of supplementary motor area. *Science* 343, 888–891

7. Hardwick, R.M. *et al.* (2013) A quantitative meta-analysis and review of motor learning in the human brain. *NeuroImage* 67, 283–297

8. Swann, N.C. *et al.* (2012) Roles for the pre-supplementary motor area and the right inferior frontal gyrus in stopping action: electrophysiological responses and functional and structural connectivity. *NeuroImage* 59, 2860–2870

9. Nachev, P. *et al.* (2008) Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* 9, 856–869

10. Rizzolatti, G. *et al.* (2014) Cortical mechanisms underlying the organization of goal-directed actions and mirror neuron-based action understanding. *Physiol. Rev.* 94, 655–706

11. Picard, N. and Strick, P.L. (1996) Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex* 6, 342–353

12. Bestelmeyer, P.E.G. *et al.* (2014) Adaptation to vocal expressions reveals multistep perception of auditory emotion. *J. Neurosci.* 34, 8098–8105

Outstanding Questions

Are there hemispheric asymmetries in auditory processing in SMA and pre-SMA? Does this differ depending on the nature of the sound, and on whether it is heard or imagined?

Does the spatial overlap of activity across auditory perception, imagery, and sound production in SMA and pre-SMA reflect the engagement of the same neural populations and mechanisms? Do these regions contain both mechanisms specialized for action as compared to perception (and vice versa) and mechanisms shared across modalities?

Do responses in SMA and pre-SMA during auditory processing and imagery correlate with the ability to produce the corresponding sounds?

How do SMA and pre-SMA interact with the nodes of the anteroventral and posterodorsal streams of auditory processing in typical as well as in atypical conditions, such as auditory verbal hallucinations?

How does damage or stimulation of SMA and pre-SMA affect auditory processing and imagery? Are these possible effects modulated by task conditions, sound category, and by the level of degradation of the sensory input?

13. Herholz, S.C. *et al.* (2012) Neuronal correlates of perception, imagery, and memory for familiar tunes. *J. Cogn. Neurosci.* 24, 1382–1397

14. Warren, J.W. *et al.* (2006) Positive emotions preferentially engage an auditory-motor 'mirror' system. *J. Neurosci.* 26, 13067–13075

15. Scott, S.K. *et al.* (2004) A positron emission tomography study of the neural basis of informational and energetic masking effects in speech perception. *J. Acoust. Soc. Am.* 115, 813–821

16. McGettigan, C. *et al.* (2015) Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cereb. Cortex* 25, 246–257

17. Silva-Pereira, C. *et al.* (2011) Music and emotions in the brain: familiarity matters. *PLoS ONE* 6, e27241

18. Vogt, C. and Vogt, O. (1919) Allgemeine ergebnisse unserer hirnforschung. *Journal für Psychologie und Neurologie* 25, 277–462

19. Zilles, K. *et al.* (1995) Mapping of human and macaque sensori-motor areas by integrating architectonic, transmitter receptor, MRI and PET data. *J Anat.* 187, 515–537

20. Luppino, G. *et al.* (1993) Corticocortical connections of area F3 (SMA-Proper) and area F6 (pre-SMA) in the macaque monkey. *J. Comp. Neurol.* 338, 114–140

21. Geyer, S. *et al.* (1998) Receptor autoradiographic mapping of the mesial motor and premotor cortex of the macaque monkey. *J. Comp. Neurol.* 397, 231–250

22. Johansen-Berg, H. *et al.* (2004) Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proc. Natl. Acad. Sci. USA* 101, 13335–13340

23. Dum, R.P. and Strick, P.L. (1991) The origin of corticospinal projections from the premotor areas in the frontal lobe. *J Neurosci.* 11, 667–689

24. Russo, G.S. *et al.* (2002) Neural activity in monkey dorsal and ventral cingulate motor areas: comparison with the supplementary motor area. *J. Neurophysiol.* 88, 2612–2629

25. Vorobiev, V. *et al.* (1998) Parcellation of human mesial area 6: cytoarchitectonic evidence for three separate areas. *Eur. J. Neurosci.* 10, 2199–2203

26. Mayka, M.A. *et al.* (2006) Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. *NeuroImage* 31, 1453–1474

27. Tanaka, S. *et al.* (2005) Modality-specific cognitive function of medial and lateral human Brodmann area 6. *J. Neurosci.* 25, 496–501

28. Muakkassa, K.F. and Strick, P.L. (1979) Frontal lobe inputs to primate motor cortex: evidence for four somatotopically organized 'premotor' areas. *Brain Res.* 177, 176–182

29. Barbas, H. and Pandya, D.N. (1987) Architecture and frontal cortical connections of the premotor cortex (area 6) in the rhesus monkey. *J. Comp. Neurol.* 256, 211–228

30. Schubotz, R.I. and von Cramon, D.Y. (2003) Functional–anatomical concepts of human premotor cortex: evidence from fMRI and PET studies. *NeuroImage* 20, S120–S131

31. Fontaine, D. *et al.* (2002) Somatopy of the supplementary motor area: evidence from correlation of the extent of surgical resection with the clinical patterns of deficit. *Neurosurgery* 50, 297–305

32. Fried, I. *et al.* (1991) Functional organization of human supplementary motor cortex studied by electrical stimulation. *J. Neurosci.* 11, 3656–3666

33. Mitz, A.R. and Wise, S.P. (1987) The somatotopic organization of the supplementary motor area: intracortical microstimulation mapping. *J. Neurosci.* 7, 1010–1021

34. Luppino, G. *et al.* (1991) Multiple representations of body movements in mesial area 6 and the adjacent cingulate cortex: an intracortical microstimulation study in the macaque monkey. *J. Com. Neurol.* 311, 463–482

35. Chainay, H. *et al.* (2004) Foot, face and hand representation in the human supplementary motor area. *NeuroReport* 15, 765–769

36. Zeharia, N. *et al.* (2012) Negative blood oxygenation level dependent homunculus and somatotopic information in primary motor cortex and supplementary motor area. *Proc. Natl. Acad. Sci. USA* 109, 18565–18570

37. Godschalk, M. *et al.* (1995) Somatotopy of monkey premotor cortex examined with microstimulation. *Neurosci. Res.* 23, 269–279

38. Sakai, K. *et al.* (1999) Presupplementary motor area activation during sequence learning reflects visuo-motor association. *J. Neurosci.* 19, 1–6

39. Dum, R.P. and Strick, P.L. (1996) Spinal cord terminations of the medial wall motor areas in macaque monkeys. *J. Neurosci.* 16, 6513–6525

40. Vergani, F. *et al.* (2015) White matter connections of the supplementary motor area in humans. *J. Neurol. Neurosurg. Psychiatry* 85, 1377–1385

41. Lu, M. *et al.* (1994) Interconnections between the prefrontal cortex and the premotor areas in the frontal lobe. *J. Comp. Neurol.* 341, 375–392

42. Akkal, D. *et al.* (2007) Supplementary motor area and presupplementary motor area: targets of basal ganglia and cerebellar output. *J. Neurosci.* 27, 10659–10673

43. Lehéricy, S. *et al.* (2004) 3-D diffusion tensor axonal tracking shows distinct SMA and Pre-SMA projections to the human striatum. *Cereb. Cortex* 14, 1302–1309

44. Catani, M. *et al.* (2012) Short frontal lobe connections of the human brain. *Cortex* 48, 273–291

45. Catani, M. *et al.* (2013) A novel frontal pathway underlies verbal fluency in primary progressive aphasia. *Brain* 136, 2619–2628

46. Luppino, G. *et al.* (2001) Projections from the superior temporal sulcus to the agranular frontal cortex in the macaque. *Eur. J. Neurosci.* 14, 1035–1040

47. Reznik, D. *et al.* (2015) Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. *J. Neurosci.* 35, 2173–2180

48. Pinho, A.L. *et al.* (2014) Connecting to create: expertise in musical improvisation is associated with increased functional connectivity between premotor and prefrontal areas. *J. Neurosci.* 34, 6156–6163

49. Jürgens, U. (1984) The efferent and afferent connections of the supplementary motor area. *Brain Res.* 300, 63–81

50. Liebenthal, E. *et al.* (2013) Neural dynamics of phonological processing in the dorsal auditory stream. *J. Neurosci.* 33, 15414–15424

51. Lee, Y. *et al.* (2012) Categorical speech processing in Broca's area: an fMRI study using multivariate pattern-based analysis. *J. Neurosci.* 32, 3942–3948

52. Binder, J.R. *et al.* (2008) A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia* 49, 1980–1997

53. LoCasto, P.C. *et al.* (2004) An fMRI investigation of speech and tone segmentation. *J. Cogn. Neurosci.* 16, 1612–1624

54. Jardri, R. *et al.* (2007) Self awareness and speech processing: an fMRI study. *NeuroImage* 35, 1645–1653

55. Peelle, J.E. *et al.* (2010) Neural processing during older adults' comprehension of spoken sentences: age differences in resource allocation and connectivity. *Cereb. Cortex* 20, 773–782

56. Gauvin, H.S. *et al.* (2016) Conflict monitoring in speech processing: an fMRI study of error detection in speech production and perception. *NeuroImage* 126, 96–105

57. Tremblay, P. and Small, S.L. (2011) From language comprehension to action understanding and back again. *Cereb. Cortex* 21, 1166–1177

58. Desai, R.H. *et al.* (2010) Activation of sensory-motor areas in sentence comprehension. *Cereb. Cortex* 20, 468–478

59. Meltzer, J.A. *et al.* (2010) Neural aspects of sentence comprehension: syntactic complexity, reversibility, and reanalysis. *Cereb. Cortex* 20, 1853–1864

60. Du, Y. *et al.* (2015) Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proc. Natl. Acad. Sci. USA* 111, 7126–7131

61. Adank, P. (2012) The neural bases of difficult speech comprehension and speech production: Two Activation Likelihood Estimation (ALE) meta-analysis. *Brain Lang.* 122, 42–54

62. Chen, J.L. *et al.* (2012) Learning to play a melody: an fMRI study examining the formation of auditory-motor associations. *NeuroImage* 59, 1200–1208

63. Brown, S. and Martinez, M.J. (2007) Activation of premotor vocal areas during musical discrimination. *Brain Cogn.* 63, 59–69

64. Gazzola, V. *et al.* (2006) Empathy and the somatotopic auditory mirror system in humans. *Curr. Biol.* 16, 1824–1829

65. Jäncke, L. *et al.* (2003) Focused attention in a simple dichotic listening task: an fMRI experiment. *Cogn. Brain Res.* 16, 257–266

66. Nastase, S. *et al.* (2014) Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. *Hum. Brain Mapp.* 35, 1111–1128

67. Sammler, D. *et al.* (2015) Dorsal and ventral pathways for prosody. *Curr. Biol.* 25, 3079–3085

68. Shergill, S.S. *et al.* (2001) A functional study of auditory verbal imagery. *Psychol. Med.* 31, 241–253

69. Tian, X. *et al.* (2016) Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex* 77, 1–12

70. Leaver, A.M. *et al.* (2009) Brain activation during anticipation of sound sequences. *J. Neurosci.* 29, 2477–2485

71. McNorgan, C. (2012) A meta-analytic review of multisensory imagery identifies the neural correlates of modality-specific and modality-general imagery. *Front. Hum. Neurosci.* 6, 285

72. Halpern, A.R. *et al.* (2004) Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia* 42, 1281–1292

73. Zatorre, R.J. *et al.* (1996) Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8, 29–46

74. Raij, T.T. and Riekki, T.J.J. (2012) Poor supplementary motor area activation differentiates auditory verbal hallucination from imagining the hallucination. *NeuroImage Clin.* 1, 75–80

75. Linden, D.E.J. *et al.* (2011) The brain's voices: comparing non-clinical auditory hallucinations and imagery. *Cereb. Cortex* 21, 330–337

76. Chen, J.L. *et al.* (2008) Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* 18, 2844–2854

77. Peretz, I. *et al.* (2009) Musical lexical networks: the cortical organization of music recognition. *Ann. N. Y. Acad. Sci.* 1169, 256–265

78. Adank, P. *et al.* (2013) The role of accent imitation in sensorimotor integration during processing of intelligible speech. *Front. Hum. Neurosci.* 7, 634

79. Margulis, E.H. *et al.* (2009) Selective neurophysiologic responses to music in instrumentalists with different listening biographies. *Hum. Brain Mapp.* 30, 267–275

80. Lima, C.F. *et al.* (2015) Feel the noise: relating individual differences in auditory imagery to the structure and function of sensorimotor systems. *Cereb. Cortex* 25, 4638–4650

81. Zvyagintsev, M. *et al.* (2013) Brain networks underlying mental imagery of auditory and visual information. *Eur. J. Neurosci.* 9, 1421–1434

82. Herrmann, B. *et al.* (2014) Supplementary motor area activations predict individual differences in temporal-change sensitivity and its illusory distortions. *NeuroImage* 101, 370–379

83. Mottonen, R. *et al.* (2013) Auditory-motor processing of speech sounds. *Cereb Cortex* 23, 1190–1197

84. Banissy, M. *et al.* (2010) Suppressing sensorimotor activity modulates the discrimination of auditory emotions but not speaker identity. *J. Neurosci.* 30, 13552–13557

85. Adank, P. and Devlin, J. (2010) On-line plasticity in spoken sentence comprehension: adapting to time-compressed speech. *NeuroImage* 49, 1124–1132

86. Kirzinger, A. and Jürgens, U. (1982) Cortical lesion effects and vocalization in the squirrel monkey. *Brain Res.* 233, 299–315

87. Bengtsson, S.L. *et al.* (2009) Listening to rhythms activate motor and premotor cortices. *Cortex* 45, 62–71

88. Grahn, J.A. and Brett, M. (2007) Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* 19, 893–906

89. Chapin, H.L. *et al.* (2010) Neural responses to complex auditory rhythms: the role of attending. *Front. Psychol.* 1, 224

90. Janata, P. and Grafton, S. (2003) Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat. Rev. Neurosci.* 6, 682–687

91. Zatorre, R.J. *et al.* (2007) When the brain plays music: auditory–motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558

92. Kotz, S.A. and Schwartze, M. (2010) Cortical speech processing unplugged: a timely subcortico-cortical framework. *Trends Cogn. Sci.* 14, 392–399

93. Schwartze, M. *et al.* (2012) Functional dissociation of pre-SMA and SMA-proper in temporal processing. *NeuroImage* 60, 290–298

94. Shahin, A.J. *et al.* (2009) Neural mechanisms for illusory filling-in of degraded speech. *NeuroImage* 44, 1133–1143

95. Haggard, P. (2008) Human volition: towards a neuroscience of will. *Nat. Rev. Neurosci.* 9, 934–946

96. Ridderinkhof, K.R. *et al.* (2004) The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447

97. Alario, F-X. *et al.* (2006) The role of the supplementary motor area (SMA) in word production. *Cortex* 1076, 129–143

98. Blank, S.C. *et al.* (2002) Speech production: Wernicke, Broca and beyond. *Brain* 125, 1829–1838

99. Tourville, J.A. *et al.* (2008) Neural mechanisms underlying auditory feedback control of speech. *NeuroImage* 39, 1429–1443

100. Christoffels, I.K. *et al.* (2007) Neural correlates of verbal feedback processing: an fMRI study employing overt speech. *Hum. Brain Mapp.* 28, 868–879

101. Benoit, C. *et al.* (2014) Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. *Front. Hum. Neurosci.* 8, 494

102. Kurata, K. *et al.* (2000) Activation of the dorsal premotor cortex and pre-supplementary motor area of humans during an auditory conditional motor task. *J. Neurophysiol.* 84, 1667–1672

103. Kasahara, K. *et al.* (2015) Neuroanatomical correlates of brain–computer interface performance. *NeuroImage* 110, 95–100

104. Ehrsson, H.H. *et al.* (2003) Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *J. Neurophysiol.* 90, 3304–3316

105. Sharma, N. and Baron, J. (2013) Does motor imagery share neural networks with executed movement: a multivariate fMRI analysis. *Front. Hum. Neurosci.* 7, 564

106. Zentner, M. and Eerola, T. (2010) Rhythmic engagement with music in infancy. *Proc. Natl. Acad. Sci. USA* 107, 5768–5773

107. Scott, S.K. *et al.* (2014) The social life of laughter. *Trends Cogn. Sci.* 18, 618–620

108. Rochet-Capellan, A. and Fuchs, S. (2014) Take a breath and take the turn: how breathing meets turns in spontaneous dialogue. *Philos. T. Roy. Soc. B* 369, 1–10

109. Scott, S.K. *et al.* (2009) A little more conversation, a little less action–candidate roles for the motor cortex in speech perception. *Nat. Rev. Neurosci.* 10, 295–302

110. Adank, P. *et al.* (2013) Accent imitation positively affects language attitudes. *Front. Psychol.* 4, 280

111. Fujioka, T. *et al.* (2012) Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *J. Neurosci.* 32, 1791–1802

112. Grush, R. (2004) The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–442

113. Schubotz, R.I. (2007) Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* 11, 211–218

114. Pfordresher, P.Q. and Halpern, A.R. (2013) Auditory imagery and the poor-pitch singer. *Psychon. Bull. Rev.* 20, 747–753

115. Navarro-Cebrian, A. and Janata, P. (2010) Electrophysiological correlates of accurate mental image formation in auditory perception and imagery tasks. *Brain Res.* 1342, 39–54

116. Morillon, B. *et al.* (2015) Predictive motor control of sensory dynamics in auditory active sensing. *Curr. Opinion. Neurobiol.* 31, 230–238

117. Wolfensteller, U. *et al.* (2007) Understanding non-biological dynamics with your own premotor system. *NeuroImage* 36, T33–T43

118. Markov, N.T. *et al.* (2014) A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* 24, 17–36

119. Scott, S.K. and Johnsrude, I.S. (2003) The neuroanatomical and functional organization of speech perception. *Trends Neurosci.* 26, 100–107

120. Rauschecker, J.P. and Scott, S.K. (2009) Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* 12, 718–724

121. Rauschecker, J. (2011) An expanded role for the dorsal auditory pathway in sensorimotor control and integration. *Hear. Res.* 271, 16–25

122. Obleser, J. *et al.* (2007) Multiple stages of auditory speech perception reflected in event-related fMRI. *Cereb. Cortex* 17, 2251–2257

123. Patterson, K. *et al.* (2007) Where do you know what you know? The representation of semantic knowledge in the human brain. *Nat. Rev. Neurosci.* 8, 976–987

124. Wise, R.J.S. *et al.* (2001) Separate neural subsystems within 'Wernicke's area'. *Brain* 124, 83–95

125. Agnew, Z.K. *et al.* (2013) Articulatory movements modulate auditory responses to speech. *NeuroImage* 73, 191–199

126. Agnew, Z.K. *et al.* (2011) Discriminating between auditory and motor cortical responses to speech and nonspeech mouth sounds. *J. Cogn. Neurosci.* 23, 4038–4047

127. Repp, B.H. (2001) Phase correction, phase resetting, and phase shits after subliminal timing perturbations in sensorimotor synchronization. *J. Exp. Psychol. - Hum. Percep. Perform.* 27, 600–621

128. Taubert, M. *et al.* (2010) Dynamic properties of human brain structure: learning-related changes in cortical areas and associated fiber connections. *J. Neurosci.* 30, 11670–11677

129. Taubert, M. *et al.* (2010) Long-term effects of motor training on resting-state networks and underlying brain structure. *NeuroImage* 57, 1492–1498

130. Hänggi, J. *et al.* (2010) Structural neuroplasticity in the sensorimotor network of professional female ballet dancers. *Hum. Brain Mapp.* 31, 1196–1206

131. Hänggi, J. *et al.* (2015) Structural brain correlates associated with professional handball playing. *PLoS ONE* 10, e0124222

132. Huang, R. *et al.* (2015) Long-term intensive training induced brain structural changes in world class gymnasts. *Brain Struct. Funct.* 220, 625–644

133. Gaser, C. and Schlaug, G. (2003) Brain structures differ between musicians and non-musicians. *J. Neurosci.* 23, 9240–9245

134. Dayan, E. and Cohen, L.G. (2011) Neuroplasticity subserving motor skill learning. *Neuron* 72, 443–454

135. Kleber, B. *et al.* (2010) The brain of opera singers: experience-dependent changes in functional activation. *Cereb. Cortex* 20, 1144–1152

136. Meister, I. *et al.* (2005) Effects of long-term practice and task complexity in musicians and nonmusicians performing simple and complex motor tasks: implications for cortical motor organization. *Hum. Brain Mapp.* 25, 345–352

137. Landau, S.M. and D'Esposito, M. (2006) Sequence learning in pianists and nonpianists: an fMRI study of motor expertise. *Cogn. Affect. Behav. Neurosci.* 6, 246–259

138. Grahn, J.A. and McAuley, J.D. (2009) Neural bases of individual differences in beat perception. *NeuroImage* 47, 1894–1903

139. Woods, E.A. *et al.* (2014) Expert athletes activate somatosensory and motor planning regions of the brain when passively listening to familiar sports sounds. *Brain Cogn.* 87, 122–133

140. Aziz-Zadeh, L. *et al.* (2010) Common premotor regions for the perception and production of prosody and correlations with empathy and prosodic ability. *PLoS ONE* 5, e8759

141. Drolet, M. *et al.* (2014) Recognizing the authenticity of emotional expressions: F0 contour matters when you need to know. *Front Hum. Neurosci.* 8, 144

142. Kreifelts, B. *et al.* (2013) Non-verbal emotion communication training induces specific changes in brain function and structure. *Front. Hum. Neurosci.* 7, 648

143. Herbet, G. *et al.* (2014) Inferring a dual-stream model of mentalizing from associative white matter fibres disconnection. *Brain* 137, 944–959

144. Rodigari, A. and Oliveri, M. (2014) Disrupting SMA activity modulates explicit and implicit emotional responses: an rTMS study. *Neurosci. Lett.* 579, 30–34

145. Farrugia, N. *et al.* (2015) Tunes stuck in your brain: the frequency and affective evaluation of involuntary musical imagery correlate with cortical structure. *Conscious Cogn.* 35, 66–77

146. Caspers, S. *et al.* (2010) ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage* 50, 1148–1167

147. Gazzola, V. and Keysers, C. (2009) The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. *Cereb. Cortex* 19, 1239–1255

148. Mukamel, R. *et al.* (2010) Single-neuron responses in human during execution and observation of actions. *Curr. Biol.* 20, 750–756

149. Bangert, M. *et al.* (2006) Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *NeuroImage* 30, 917–926

150. Van de Ven, V. *et al.* (2009) Neural network of speech monitoring overlaps with overt speech production and comprehension networks: a sequential spatial and temporal ICA study. *NeuroImage* 47, 1982–1991

151. Caramazza, A. *et al.* (2014) Embodied cognition and mirror neurons: a critical assessment. *Annu. Rev. Neurosci.* 37, 1–15