How passive is passive listening?
Toward a sensorimotor theory of auditory perception

Tom Froese[1,2] · Ximena González-Grandón[1,3]

Abstract

According to sensorimotor theory perceiving is a bodily skill involving exercise of an implicit know-how of the systematic ways that sensations change as a result of potential movements, that is, of sensorimotor contingencies. The theory has been most successfully applied to vision and touch, while perceptual modalities that rely less on overt exploration of the environment have not received as much attention. In addition, most research has focused on philosophically grounding the theory and on psychologically elucidating sensorimotor laws, but the theory's ramifications for neuroscience still remain underexamined. Here we sketch the beginnings of a research program that could address these two outstanding challenges in terms of auditory perception. We review the neuroscience literature on passive listening, which is defined as listening without overt bodily movement, and conclude that sensorimotor theory provides a unique perspective on the consistent finding of motor system activation. In contrast to competing theories, this activation is predicted to be involved not only in the perception of speech- and action-related sounds, but in auditory perception in general. More specifically, we propose that the auditory processing associated with supplementary motor areas forms part of the neural basis of the exercise of sensorimotor know-how: these areas' recognized role in (1) facilitating spontaneous motor responses to sound and (2) supporting flexible engagement of sensorimotor processes to guide auditory experience and enable auditory imagery, can be understood in terms of two key characteristics of sensorimotor interaction, its (1) "alerting capacity" (or "grabbiness") and (2) "corporality" (or "bodiliness"), respectively. We also highlight that there is more to the inside of the body than the brain: there is an opportunity to develop sensorimotor theory into new directions in terms of the still poorly understood active processes of the peripheral auditory system.

Keywords Auditory perception · Enactivism · Bodily skill · Sensorimotor theory · Motor system · Dual-stream model · Motor resonance theory

✉ Tom Froese
t.froese@gmail.com

Extended author information available on the last page of the article

1 Introduction

Action-dependent characteristics of neural activity related to auditory perception have gained increasing attention in the last couple of decades. Neurophysiological and neuroimaging research has shown that brain areas classically thought to be involved only in action production can also be recruited in auditory perception of actions, especially of speech, music, and tool-use, but also more generally of environmental sounds (e.g. D'Ausilio et al. 2014; Fadiga et al. 2002; Meister et al. 2003; Pulvermüller et al. 2006; Watkins et al. 2003). The involvement of these brain areas cannot simply be explained in terms of the perceiver's movements: auditory perception is associated with activation of the brain's motor areas even during passive listening, that is, even while the listener does not engage in overt movements (Lima et al. 2016).

This motor system activation is rather unexpected from the perspective of a traditional theory of perception, which treats perceiving as an essentially passive process of sensory information processing in the service of building detailed internal mental representations of the external environment. Instead, this evidence lends itself to an interpretation in terms of a sensorimotor theory of perception. As an illustrative example, consider these findings of a study of speech perception:

> In sum, this study confirms the central role of bilateral superior temporal regions in speech perception, since only in these areas did signal change correlate with the producibility of novel phonemes. However, there is also evidence for the involvement of motor areas in speech perception, as speech motor areas were activated by passive speech perception, distinguished between native and non-native phonemes, and were functionally connected with superior temporal cortex. Taken together, these findings constitute evidence for the sensorimotor nature of speech perception. (Wilson and Iacoboni 2006: 323)

However, this tantalizing potential connection between evidence of motor system activation and sensorimotor theory is still in need of further development, both because the sensorimotor theory of auditory perception requires clarification and because its neuroscientific predictions must also still be made more explicit. Moreover, there are other theories of auditory perception that can also account for this motor system activation, and yet they do not agree on what, if any, function this activation has in perception. For instance, the classical dual stream model can interpret motor system activation as a largely nonfunctional result of neural associations formed by experiencing repeated action-perception co-occurrences. On the other hand, the broad evidence for motor activation suggests that these regions need to be conceptualized as part of the functional neuroanatomy of auditory processing (Lima et al. 2016). For example, a recent dual stream model proposes that it serves the function of an internal forward model that predicts upcoming sensory events (Bornkessel-Schlesewsky et al. 2015), while the motor theory of speech perception and motor resonance theory more broadly propose a role for motor simulation of the other's acoustic gestures, specifically in perceiving speech- and action-related sounds, respectively (Galantucci et al. 2006).

However, these other theories' frequent appeals to mental content-carrying functions of this motor activity sits uneasily with sensorimotor theory's preference for accounting of perception in terms of direct access to the environment (Beaton

2016). In contrast, the sensorimotor theory claims that perceiving is a bodily skill involving exercise of know-how of the systematic ways that sensory inputs are disposed to change as a result of potential movements, and this bodily skill is best conceived of in non-representational terms (Silverman 2018), and as bringing sensory neural activity under the influence of potential interactions with the environment (Degenaar and O'Regan 2015). We therefore propose that the notion of exercising sensorimotor know-how predicts a necessary functional role of motor activation in all forms of auditory perception, including in passive listening to sounds that are both novel and not directly controllable by others' gestures. If this were to be empirically confirmed it would at the same time pose challenges to theories based on neural association and simulation, respectively.

In the following section we first review the sensorimotor theory of perception in more detail and then consider the applicability of its core commitments to the auditory modality. We then compare and contrast three kinds of theories of auditory perception, the dual stream model, motor resonance theory, and sensorimotor theory, especially with respect to the divergent predictions they make regarding possible functions of motor system activation. This is followed by an evaluation of the current evidence of motor system activation during passive listening, and we conclude with a discussion of how these theories' divergent predictions fare with respect to this evidence.

2 Sensorimotor theory of perceptual experience

The enactive approach to perception, and to cognitive science more broadly, was launched nearly three decades ago with the claim that "perception consists in perceptually guided action" leading to an overall concern "to determine the common principles or lawful linkages between sensory and motor systems" (Varela et al. 1991: 173). This overall concern was a key inspiration for a distinctive philosophical and scientific research program that has been principally developed by O'Regan and colleagues (e.g. Noë 2004; O'Regan and Noë 2001; O'Regan 2011), and which has come to be known as sensorimotor enactivism (Ward et al. 2017) or as sensorimotor theory (Bishop and Martin 2014).

Sensorimotor theory and the enactive approach both reject mainstream theories of perception, which claim that perceiving is essentially about generating internal mental representations of the external world. This rejection of internal representation leads to two broad possibilities: on the one hand, the basis of perception remains internal to the brain, but the process is instead conceived of in constructivist rather than representational terms; or, on the other hand, the basis of perception extends beyond the brain and into the sensorimotor interactions, and accordingly the process can provide direct access to the external world (Vörös et al. 2016). Sensorimotor theory favors the latter possibility and can thus be distinguished from some earlier formulations of the enactive approach to perception, whose constructivist sympathies can be criticized for leading to a lingering neuro-centric internalism about the constitutive basis of experience (Beaton 2013; Pascal and O'Regan 2008). However, most recent formulations of enactivism explicitly defend the claim that the basis of the experience is not just in the head (Di Paolo et al. 2017; Fuchs 2018; Hutto and Myin 2013), and so this particular difference no longer applies.

A genuine difference that remains concerns the question whether perceptual skills require a form of agency grounded in biological embodiment: while the enactive approach emphasizes the strong continuity between life and mind (Di Paolo et al. 2017; Kirchhoff and Froese 2017; Noë 2009; Thompson 2007), and therefore is skeptical about the possibility that robotic agents have perceptual experience (Froese and Ziemke 2009), sensorimotor theory is more open to this possibility (Degenaar and O'Regan 2017). This difference, however, is not directly relevant to the sections that follow, and we assume that what we have to say about sensorimotor theory applies to all of these variants of enactivism.

Famously, sensorimotor theory is not only concerned with providing an account of perception; it also has set itself the much more challenging task of explaining sensory consciousness:

> We suggested that "feel" should be construed, not as a brain activity, but as a skill that the organism exercises. The skill we refer to involved having mastery of what we called the sensorimotor contingencies – that is, the laws of co-variation between actions of the organism and resulting changes in sensory input. (O'Regan et al. 2005: 384)

For newcomers sensorimotor theory is typically motivated in an intuitive manner by highlighting that perceiving the material properties of sponge requires actively manipulating the sponge with one's hand, a process which extends in time (Silverman 2013). And yet this example involving overt movement can lead to confusion if it is taken too literally (Aizawa 2019), and must therefore be immediately followed by the caveat that such overt movement is normally not necessary for the constitution of perceptual experience (Myin 2016). For instance, as long as the perceiver has a history in handling sponges or similar objects, it is sufficient for her to passively hold the sponge to have a perceptual experience of its material properties. In other words, perceptual experience is claimed to be constituted by the bodily skill of knowing how tactile sensations *would* change as a result of potential overt hand movements, and so it does not require *actual* change, although effecting actual change by self-movement has the effect of improving the veracity of perceptual experience by confirming the anticipated sensorimotor regularities.

Indeed, if perceptual experience is constituted by potential exploratory movements it may turn out to be misleading, which has been amply demonstrated in the case of the visual modality in terms of change blindness and inattentional blindness. These experimental paradigms have revealed that we can fail to see prominent changes in the world even though we have the experience of seeing a richly detailed world (Chabris and Simons 2010; O'Regan et al. 1999). Sensorimotor theory holds that implicit know-how constitutes this experienced richness in terms of the perceptual accessibility of the currently non-accessed detail. The same argument also applies to the perceptual presence of objects within the world (Noë 2002): I perceive a tomato as a complete three-dimensional object, rather than as just a façade without a back, even though I currently only have visual access to its front. My experience of the visually absent sides of the tomato as being implicitly co-present with the visually present front is constituted by my sensorimotor know-how of how the presently occluded sides would systematically become visually present were I to move around the tomato or to

turn the tomato around. The phenomenology of perceptual richness and perceptual presence both provide powerful reasons for thinking that perceptual experience is constituted by bodily skills (Silverman 2018).

Sensorimotor theory is particularly compelling for the visual and tactile modalities, both of which are inherently exploratory because they are characterized by narrow receptor organs (e.g. the retina and fingers tips) and subject to occlusion (e.g. I see or feel the whole sponge even if I cannot directly sense all of the sides simultaneously). These anatomical and structural limitations arguably play a fundamental role in accounting for the fact that potential bodily movement is essential for vision and touch.[1] However, more work needs to be done to show that these arguments in favor or sensorimotor theory generalize to other perceptual modalities (Lyon 2014), especially if they are not as constrained in this same manner, like olfaction (e.g. Cooke and Myin 2011) and auditory perception (e.g. Aytekin et al. 2008). Here we focus on developing sensorimotor theory with respect to the auditory modality.

2.1 Toward a sensorimotor theory of auditory experience

Generalizing sensorimotor theory from vision and touch to audition implies that auditory perceiving should be understood as a bodily skill, specifically as involving mastery of the corresponding sensorimotor contingencies, that is, of the laws of co-variation between movements of the perceiver and resulting changes in auditory sensations. O'Regan et al. (2005) analyzed this co-variation in terms of its "corporality" and "alerting capacity", which are properties that correspond to the perceiver's specific exploratory relations with the environment. They help to account for the fact that perceptual experiences have a special phenomenal quality that is not shared by other mental phenomena, and which is specific to the different perceptual modalities, including the auditory modality.

Here are some examples of the *corporality* of auditory sensorimotor contingencies:

"They are affected in special ways by head movements: rotations of the head generally change the temporal asynchrony between left and right ears. Movement of the head in the direction of the sound source mainly affects the amplitude but not the frequency of the sensory input." (O'Regan and Noë 2001: 941)

"turning your head immediately affects the phase and amplitude difference between signals coming into the two ears; approaching a sound source increases the amplitude" (O'Regan et al. 2005: 374)

"sound gets louder as you move toward its source but stays constant in volume when you travel at a fixed distance around it" (Di Paolo et al. 2017: 18)

---

[1] In the case of vision this potential access can be interpreted as a saccade followed by fixation on the relevant visual detail. However, this is not the full story as change blindness can also occur when the eye is directly fixating the change at the moment it occurs, suggesting that "you do not always see where we look" (O'Regan et al. 2000: 201). In other words, it is important to keep in mind that accessing detail with the eyes is not sufficient in itself for visual awareness, which additionally requires attention.

"There are constancy effects as well in the domain of hearing. The siren sounds louder as we approach it, even though we can also tell, by hearing, that it hasn't actually gotten louder. We hear it as louder thanks to the change in our relationship to it. Likewise, you can perceive that the person near you and the person over there are talking at about the same volume, even though the person nearer prevents you from hearing the one over there (the sound drowns out the other, which is auditory occlusion). How things sound varies as one moves in relation to sound sources. The pattern of change as one moves make the world available to perception." (Noë 2004: 161)

And here are some examples related to the *alerting capacity* of sounds:

"In the auditory system there are head-orienting reflexes which cause you to turn to sudden noises" (O'Regan et al. 2005: 375).

"A sharp sound makes us turn in the direction from which the sound emanates. […] A person speaks to us, we turn to him or her. In this sort of way, and in countless ways like this, sensory impressions are immediately coupled with spontaneous movement." (Noë 2004: 6)

"We continuously adjust our head in order to better take in noises belonging to one happening or another, and the fact that we do so shows that we are implicitly familiar with the way changes in our relation to events alters our sensory stimulation." (Noë 2004: 160)

The corporality and alerting capacity of auditory experience strongly predict that potential bodily movements play a functional role in auditory processing even during passive listening tasks, specifically in aiding the perceiver to make sense of sound events and to potentially respond to them, respectively. We can already note that this prediction is at least broadly consistent with the current neuroscience literature, which is converging on the hypothesis that consistent activation of supplementary motor areas during passive listening might have two distinct functions, namely to "guide perception" and to "facilitate behavioural responses to sounds" (Lima et al. 2016: 527). However, sensorimotor theory goes further by claiming that implicit know-how of how sounds are disposed to change in terms of potential bodily movements not only guides perception, but is necessary for it because it actually *constitutes* perceptual experience. For the visual modality this stronger, constitutive claim is typically motivated by using it to provide accounts of perceptual richness and perceptual presence, so it is important to consider whether these accounts can be generalized to the auditory modality.

2.2 Toward a sensorimotor theory of auditory richness

We certainly have the perceptual experience of being situated in a meaningful soundscape that is richly detailed, and consisting of a variety of sound source categories (Brefczynski-Lewis and Lewis 2017). Anatomically, ears have access to a much broader range of stimuli compared to the narrow field of vision: sound waves come in parallel from all spatial directions around the perceiver, and they can originate from

multiple overlapping sources. Accordingly, a lot of the phenomenology of auditory richness might actually be directly constituted by the rich acoustic details that are simultaneously present to the ears, rather than by the perceiver's sensorimotor know-how of how to practically access currently non-present acoustic details. Determining how much auditory richness is constituted by acoustic stimuli compared to bodily skill in accessing it is difficult if not impossible to do in a priori manner and calls for auditory equivalents of the change blindness and inattentional blindness experimental paradigms. Although some have doubted that these paradigms can be generalized to the auditory modality (Isel 2001), current evidence suggests that it is indeed meaningful to talk about change deafness and inattentional deafness.

To some extent inattentional deafness was already demonstrated by early dichotic listening studies (Moray 1959), which have paradoxically become famous for the so-called cocktail party effect: we can reorient our attention when hearing our name mentioned in an otherwise unattended auditory stream. However, actually this reorientation happens in only about a third of participants (Wood and Cowan 1995), thus demonstrating that the majority of participants suffers from inattentional deafness even to meaningful stimuli such as their names. More recently, this inattentional deafness has been rediscovered and further substantiated: when asked to attend to conversations in a recording, participants routinely missed a character repeating the phrase "I'm a gorilla", thus replicating the famous unseen gorilla experiment in the auditory modality (Dalton and Fraenkel 2012). Other work has shown that inattentional deafness can also be induced by increased attentional load in the visual domain (Mcdonald and Lavie 2011; Raveh and Lavie 2015), and that inattentional blindness and inattentional deafness can co-occur even for noxious acoustic stimuli (Wayland et al. 2005).

However, in contrast to inattentional blindness, which is facilitated by the impossibility of the eyes fixating on a whole visual scene all at once, inattentional deafness does not involve an anatomical limitation of the ears. Both the attended and unattended acoustic stimuli are always directly present to the ears. Instead the failure of auditory perception during inattentional deafness, at least when induced by conditions of high visual load, has been associated with reduced sensory processing of auditory stimuli (Molloy et al. 2015). This difference has implications for how to think about the role of bodily movement in audition: whereas normal visual perceiving seems to require overt eye movements (cf. Aizawa 2019), or at least bodily movements with saccadic character (Gilchrist et al. 1997), normal auditory perceiving does not require overt movement at all, as is easily confirmed by the fully immersive experience afforded by using headphones to listen to a performance of your favorite band while lying still with your weight supported by your bed.

However, it seems reasonable to assume that a perceiver's capacity for movement still remains necessary for the development of normal auditory perception in the first place, for instance in terms of acquiring the skill of perceiving the spatial location of sounds in terms of the sensory consequences of overt movements (Aytekin et al. 2008). And this implies that we can account for the auditory experience of a soundscape characterized by a richly detailed spatiality as constituted by an implicit know-how of how acoustic stimuli in both ears would change with respect to potential bodily movements. Moreover, this constitutive role is presupposed by the cocktail party effect and other demonstrations of inattentional deafness that include an experienced spatial distance to the unattended auditory source.

There is also evidence for something akin to change blindness in the auditory modality, as revealed by a phenomenon known as the auditory continuity illusion (Riecke et al. 2008). While in change blindness the visual differences between two alternating images are not seen when the switch coincides with visual "mudsplashes" (O'Regan et al. 1999), in this phenomenon a sound that is normally audibly interrupted by a silent gap may be perceived as continuous when that gap is filled with another sound, such as noise. This experienced continuity has been associated with activity in the primary auditory cortex, which has been interpreted as showing that primary auditory areas play a role in representation of an auditory scene (Riecke et al. 2007).

However, other evidence suggests that this experienced stimulus continuity is not based on a mechanism that concretely fills in the missing acoustic stimuli (auditory induction), and that it rather consists in a merely apparent completeness (Repp 1992). The enactive approach to perception also questions appeals to filling-in mechanisms because it holds that the primary task of perceiving is to guide the agent's interaction with the world, not to produce internal representations (Pessoa et al. 1998). As was mentioned previously, it therefore proposes that the visual experience of a richly detailed world is constituted by mastery of sensorimotor contingencies, providing rapid potential access to the absent sensory details (Noë 2002). What potential movement-based access could refer to in the auditory domain is less clear, but it could be interpreted as a reorientation of attention, which is facilitated by potential head movements that serve to disambiguate the two interfering sounds in terms of their spatial locations. We would therefore expect motor system activation to be associated with the experienced presence of the sound during physical stimulus absence. That this tentative proposal is on the right track is suggested by an fMRI study of the auditory continuity illusion based on perception of artificially degraded words, which found that a network including the pre-Supplementary Motor Area (pre-SMA) "showed greater activity when less sensory information is available" (Shahin et al. 2009: 1139). We will later argue that the SMAs are involved in the exercise of mastery of auditory sensorimotor contingencies.

2.3 Toward a sensorimotor theory of auditory presence

The discussion of the auditory continuity illusion naturally leads to a consideration of the auditory equivalent of perceptual presence. It has long been known that we sometimes hear speech and sounds that are not currently present to the ears (Dannenbring 1976; Warren et al. 1972), which has been explained by the need to maintain stable auditory perception in our frequently noisy environment. In other words, just like vision presents us with integrated objects even though they are necessarily visually occluded by themselves (e.g. by the tomato's front side) and possibly also by other objects (e.g. another tomato partially placed in front of it), audition also presents us with integrated objects (e.g. hearing a word) even though the sequence of sounds is not all co-present at once and even though parts of the sequence can be acoustically occluded by other sounds (e.g. noise).

Importantly, both visual objects and acoustic objects are only partially accessible by sensory organs at each moment in time. As Husserl (1966) famously observed, we implicitly experience a familiar melody in its entirety even if only a short fragment of it is currently acoustically perceived. To be fair, a dynamically unfolding melody is not

directly comparable to a static apple, because the melody does not seem to require bodily movement for further aspects to be made present (in this respect it is perhaps more like seeing an apple on an automatically rotating table that automatically reveals its different sides). Yet we can assume that in both auditory and visual experience the capacity for perceptual integration and localization first required learning how sensations are disposed to change in terms of bodily movements, and that this bodily know-how is constitutive of experiencing objects as presenting more aspects than are directly given. This is not to say that perceptual learning is necessary for a subject to be sensitive to auditory stimulation as such, but it is necessary for developing that sensitivity into a modality serving meaningful auditory perception.

Moreover, as the speech versions of the auditory continuity illusion reveal, auditory presence can be about more than just temporal continuity. As Heidegger ([1927] 1996: 153) observed, we normally experience hearing meaningful objects in the world, rather than mere acoustic stimuli. For example, I hear a car hiss by my window; I normally do not hear a meaningless hiss that I then infer to be made by a car moving outside. According to sensorimotor theory, the location of the stimuli is constituted by the subject's auditory sensorimotor know-how (Aytekin et al. 2008). For example, I hear the car as present outside of my window because I implicitly know how the acoustic intensity in my ears would systematically change if I were to move to move my head with respect to the window. It is also quite likely that additional phenomenological richness is constituted by cross-modal sensorimotor know-how: if I were to move my head to look out the window in response to hearing the hiss, I would see a car speeding down the road. We will return to a consideration of multimodal perception in subsequent sections.

2.4 Key predictions of the sensorimotor theory of auditory experience

We have argued that the key tenets of sensorimotor theory can be generalized, with minor adjustments, from the visual and tactile modalities to the auditory modality. The specific auditory "feel", richness, and presence of auditory experience is constituted by mastery of distinct audio-motor contingencies, which are characterized by audio-motor versions of "corporality" and an "alerting capacity". We predict that this mastery has to be acquired during ontogenetic development on the basis of overt movement, but that expert perceivers no longer have to engage in overt movement in order to exercise their mastery. Nevertheless, we note that overt movement continues to confer small advantages even to experts, for example as demonstrated experimentally in the case of auditory localization (Perrett and Noble 1997; Rosenblum et al. 2000).

Exercise of the mastery of sensorimotor contingencies is a skill of an embodied subject acting in the world (Hurley and Noë 2003), and this organism-environment interaction process will also be crucially mediated by the nervous system (Fuchs 2018). Accordingly, we propose that from the perspective of sensorimotor theory it is expected that auditory perception activates the motor system of the brain. More specifically, motor system activation is predicted to play an essential mediating role in the perceiver's auditory experience, enabling the perceiver to exercise their bodily skill of knowing how auditory sensations are disposed to change as a result of potential movements. This prediction applies to all categories of sound stimuli. In addition, given that this bodily skill can then be cultivated by explicit auditory training,

especially in the context of rhythms and music, it is also expected that expertise in auditory perception is correlated with comparatively greater motor system activation. This is because expertise allows the perceiver to engage auditory events with a more diverse and at the same time more finely controlled set of potential movements. Finally, given that each perceptual modality has its own characteristic corporality, we should expect that motor system activation depends on the number of modalities involved. In other words, simultaneously exercising sensorimotor know-how of other perceptual modalities, in addition to audition, should lead to comparatively stronger motor system activation. We will now contrast these predictions with those of competing theories and then develop them in more detail.

3 Key predictions of competing theories of auditory perception

Sensorimotor theory is not the only theory of auditory perception that can account for motor system activation during motionless auditory perceiving. However, these other theories make divergent predictions regarding the function and scope of this activation.

3.1 Dual-stream model

An early formulation of the dual-stream model is Wernicke's proposal that dual processing streams are involved in speech perception. Following this dual-stream perspective, Rauschecker (1998) proposed the 'what–where' model: the ventral stream allows for auditory object recognition and the mapping of sound to meaning (the 'what'), while the dorsal stream enables sensorimotor mapping of the spatial nature of sound to motor and articulatory networks (the 'where'). Later Hickok and Poeppel (2007) developed an alternative dual-stream model of speech processing that maintains two separated "streams": one more closely related to speech perception and the other to speech production. The difference between these models may not be great; Rauschecker and Scott (2009) argued that the 'what/where' and 'perception/action' theories differ mainly in emphasis. Importantly, however, although Hickok and Poeppel acknowledge that the action stream indicates "a very tight connection between processes involved in speech perception and speech production," they also propose that "it does not appear to be a critical component of the speech perception process under normal (ecologically natural) listening conditions" (Hickok and Poeppel 2004: 67).

One way of interpreting the possible auditory function of areas normally associated with speech production is in terms of internal forward model structures, or "emulators" as they are known in motor control theory (Grush 2004). They involve the production of efference copies of motor commands to provide expectations of the sensory feedback and thereby enhance sensory processing (Wolpert et al. 1998). For instance, in the 'what-where' model the functions of the postero-dorsal pathway can be subsumed under a forward model: an efference copy sent from prefrontal and premotor cortices provides the basis for "optimal state estimation" in the inferior parietal lobe and in sensory areas of the posterior auditory cortex (Rauschecker 2018). In this way the motor system can be said to be involved in the prediction of sensory events, such as would be caused by the perceiver playing the next tone in a melody (Stephan et al. 2018). The purpose of such an internal forward model could also be to predict the next

action of the person that generated the sound (Gordon et al. 2018b). Nevertheless, this motor influence on perception via sensory prediction seems to be limited; for example, some evidence shows that this influence is modulatory but not necessary for speech perception (Hickok et al. 2011).

Despite its popularity, the dual-stream model has not been without criticism (Skipper et al. 2017). One concern has been the oversimplification of the architecture of the brain as a hierarchically organized structure with a "top" level formed by associative cortices (areas with increasingly complex response properties), and "low" levels formed by sensory and motor cortices. This architecture involves a division of labor between speech perception and production, involving two cortical streams of neural activity from auditory cortex to prefrontal regions (Bornkessel-Schlesewsky et al. 2015). While these streams are anatomically connected with pre-SMA and SMA, the role of these motor areas in the dual-stream model are still not well understood (Lima et al. 2016).

It is clear that the dual-stream model does not assign an essential function to motor system activation during auditory perception. Some versions of the dual-stream model accept that motor system activation can also play a functional role in audition, but this function is limited to a modulatory role based on implementing forward models that generate sensory predictions. Importantly, the scope of this predictive function is limited to the perception of auditory experiences that can be generated by bodily movements of self and others, especially speech, vocalizations, rhythms, music, and action-dependent sounds.

3.2 Motor resonance theory

Motor resonance is associated with the activation of the "mirror neuron" system, which is based on a specific class of neurons that discharge both during action production and during action perception that were first described in the monkey's brain (Rizzolatti et al. 1996). Motor resonance theory proposes that during the perception of actions performed by others, the patterns of motor system activity underlying the perceived action are mapped onto the perceiver's own motor system, resulting in congruent neural activity that serves as a motor representation of the other's action goal (Rizzolatti et al. 2001; Rizzolatti and Sinigaglia 2016). This principle of congruence is at the core of classical formulations of motor resonance theory: "A critical feature of the mirror system is the *selective* mapping of *specific* heard or seen actions onto the motor program for executing the *same* actions" (Gazzola et al. 2006: 1826).

The principle of congruence leads to the specific prediction that the functionality of the perceiver's motor system activation is at least partially determined by its similarity to what would occur if the observer himself were executing, planning, or even just imagining the perceived action. This fundamental mechanism involved in perception of action goals can be succinctly stated as follows: "individuals recognize actions made by others because the neural pattern elicited in their premotor areas during action observation is similar to that internally generated to produce that action" (Rizzolatti and Arbib 1998: 190). Many studies have confirmed that there is indeed a similarity in activation patterns. For example, an fMRI study of passive listening to sounds produced by hand- and mouth-related actions found a somatotopic pattern of activation in the left premotor cortex (Gazzola et al. 2006). An EEG study of passive listening to sounds of finger- and tongue-related actions found that these sounds evoked greater

activity in motor areas normally related with finger and tongue movement compared to control stimuli with similar physical features but without clear action associations (Hauk et al. 2006). Similarly, a study of a 12-year-old child undergoing intracranial monitoring found that the functionally defined hand area of the motor cortex exhibited activity changes selective for natural finger-clicks compared with control sounds (Lepage et al. 2010).

Emphasis on the principle of congruence makes motor resonance theory a close ally of the simulation theory of social cognition (Gallese and Goldman 1998). However, in many cases the motor system activation in the perceiver and the perceived is far from being highly similar, and so some mirror neuron researchers prefer to interpret its function in terms of neural reuse rather than as a direct mapping based on resemblance (Gallese and Sinigaglia 2018). Accordingly, motor resonance theory is not limited to predicting a strictly somatotopic organization of motor activation. Another possibility is to interpret the function of this activation not as a direct mapping of action goals, but rather in terms of the inference of intentions within a predictive coding framework (Kilner et al. 2007). Yet another source of debate in the literature is the flexibility of the mapping: possibilities range from genetically fixed correspondences to flexible associations based on lifetime learning (Cook et al. 2014). Nevertheless, most researchers agree that familiarity and expertise are important for the mirror neuron mechanism, because motor resonance requires that the observed action is within the perceiver's motor repertoire.

While most work on mirror neurons has focused on the visual modality, there has also been considerable research on the perception of actions by perceiving the sounds they produce. For example, a single-cell recording study reported a discharge of audiovisual mirror neurons located in the monkey F5c premotor area when the animal performed the action of ripping a piece of paper causing the appropriate sound, and also when it only perceived the same action-related sound (Keysers et al. 2003; Kohler et al. 2002). Evidence supporting an auditory mirror neuron system has also been found in humans (Pizzamiglio et al. 2005), including for speech perception (Fadiga et al. 2002). In this way motor resonance theory updates and extends the motor theory of speech perception (Galantucci et al. 2006; Rizzolatti and Arbib 1998), which was originally proposed by Liberman (Liberman et al. 1967; Liberman and Mattingly 1985). Motor resonance theory generalizes motor system involvement in auditory perception from speech to the whole range of action-related sounds, and frees the motor resonance mechanism from Broca's area and elevates it to a general brain principle (Rizzolatti and Sinigaglia 2016).[2]

We can derive several predictions from motor resonance theory's claim that the function of the activation of the motor system during action perception is to instantiate motor representations used for understanding the other's movements. First, it is therefore expected that impairment of the motor system is associated with impairment of action understanding, and moreover that this impairment is specific to the actions normally realized by the impaired areas. Second, it is also expected that motor system activation should function independently of the perceiver's modality, given that it is supposed to represent the other's action no matter how it is perceived. Third, given that

---

[2] In this more general formulation, motor resonance theory has substantial overlap with other important theories in addition to the motor theory of speech perception, such as grounded cognition (Barsalou 2008), common coding theory and ideomotor theory (Hommel et al. 2001), as well as motor imagery (Decety et al. 1994) and emulation theory (Grush 2004).

this function is specifically tied to action understanding, motor system activation should only be found during perception of action-related sounds that are within the perceiver's motor repertoire.

4 Evidence of motor system involvement in passive listening

In this section we evaluate the evidence of motor system involvement in auditory perception. After first introducing the passive listening experimental paradigm, we focus on three key topics that will help us to assess the predictions made by the dual stream model, motor resonance theory, and sensorimotor theory.

4.1 Passive listening tasks

The most striking evidence for motor system involvement in auditory perception comes from passive listening tasks. Passive listening tasks refer to listening to sounds without overt bodily movement (motionless listening).[3] A typical passive listening task involves presentation of auditory stimuli simultaneously to both ears via headphones and participants are instructed to listen passively to the acoustic stimuli, which includes explicit instructions to lie as still as possible with the eyes closed, and to strictly refrain from any movement, including finger lifting, and from any silent or overt vocalization, including counting (Steinbrink et al. 2009). Sometimes passive listening tasks can involve demobilization of the body, for example by fixing the head in position.

Nevertheless, in many cases eye movements, such as saccades, and head and hand micro-movements are difficult to control. There are experimental protocols where blinking is addressed (Mathias et al. 2015), and motion tracking can be used to control for finger movements (Lahav et al. 2007). Nevertheless, readers should be aware that not all studies we cite below apply such strict movement controls.

The activation of motor areas during auditory processing is by now well established (Lima et al. 2016), with most research focusing on perception of speech (Fowler and Xie 2016; Galantucci et al. 2006; Schomers and Pulvermüller 2016), of music and rhythms (Gordon et al. 2018a; Koelsch 2011; Ross et al. 2016), and of action-related sounds (Agliotti and Pazzaglia 2010, 2011). Given that each of these areas already counts with extensive reviews of the literature, we will not attempt another exhaustive review of the relevant literature here. Instead, we take a closer look at studies that speak to the functional role of motor system activation, its relationship to expertise, and its scope with respect to action-related and non-action-related sounds.

4.2 The function of motor system activation

So far there is no consensus position on what, if any, role is played by motor system activation during motionless listening (Lima et al. 2016). Indeed, despite the long tradition of research of motor resonance theory, it is still possible that much of this

---

[3] Note that this passive listening paradigm not only differs from the active listening paradigm, which is based on tasks requiring a motor response, but also from passive listening in the context of auditory selective attention task, in which passive listening is contrasted with "active" or "attentive" listening (Toro et al. 2005).

activation is an epiphenomenon, and that the "triggering of mirror neurons by action sounds represents a learned association that could, in principle, be established between an arbitrary sound and an arbitrary motor representation" (Caramazza et al. 2014: 10). Nevertheless, there are several lines of evidence that suggest that this activation does play a critical role in auditory experience. The strongest evidence comes from lesion studies, which found category-specific impairments in patients' understanding of visually perceived words (Dreyer et al. 2015) and of action-related sounds (Pazzaglia et al. 2008).

In addition, psychological studies found evidence that passive listening to a newly learned musical piece can enhance motor performance in the absence of physical practice (Lahav et al. 2005; Lahav et al. 2013), which suggests that the motor system activation typically associated with passive listening can indeed play a motor function under some circumstances. Further evidence comes from a study of performance on auditory memory, which found that musicians were more accurate at recognizing pitch alterations if they had previously performed rather than merely listened to the piece, and this was associated with greater current density in premotor and supplementary motor regions (Mathias et al. 2015), which suggests that the motor system activation plays a role in auditory discrimination. A motor-excitability study of passive listening to music found increased motor excitability for rehearsed compared to non-rehearsed pieces, led the authors to speculate that different amounts of motor activity involvement corresponds to qualitatively different states of consciousness (D'Ausilio et al. 2006). A review of motor system involvement in beat perception concluded that there is still a scarcity of research into its causal role, but that the weight of the existing studies is considerable due to the strengths of the causal designs (Ross et al. 2016).

In the context of speech perception, a TMS-study found evidence that perturbing the motor cortex had effects on speech sound discrimination, thus supporting the idea that motor structures provide a specific functional contribution to the perception of speech sounds (D'Ausilio et al. 2009). Consistent with this finding, a fMRI-study demonstrated that older adults show higher activation of frontal speech motor areas during a syllable identification task compared to young adults, and this increased activity correlates with improved speech discrimination in the older adults (Du et al. 2016). A review of passive speech perception tasks and TMS studies concluded that frontoparietal cortices, including ventral motor and somatosensory areas, are involved during speech perception and exert a causal influence on language understanding (Schomers and Pulvermüller 2016).

Taken together, the evidence converges on the side of a causal function of motor system activation for auditory perception, rather than it being merely spurious co-activation. However, given that brain areas can be recruited for different functions in different contexts, we may wonder if this causal function is related to the motor function normally performed by these brain areas during action execution, or whether they play another kind of function during perception. The most compelling evidence that the motor system activation is functional, and that this function is specifically motoric, comes from speech perception. For instance, a TMS study has demonstrated that there is a phoneme-specific increase of motor-evoked potentials related to the listener's tongue muscles when the presented words strongly involve, when pronounced, tongue movements (Fadiga et al. 2002). This motor effect is not limited to speech articulation. A TMS study that recorded motor-evoked potentials from hand and foot muscles found that these were specifically modulated by listening to hand- and

foot-action-related sentences, respectively (Buccino et al. 2005). More generally, by now it is well established that passive listening to action-related-sounds tends to modulate muscle activity in the relevant body parts (Aglioti and Pazzaglia 2010), a phenomenon known as "muscle specificity" (Gordon et al. 2018b). Moreover, this motor effect seems to play a causal function in perception. A TMS study employing ultrasound tissue Doppler imaging revealed that phoneme-specific tongue kinematic synergies were more evident in participants showing good performances in discriminating speech in noise (D'Ausilio et al. 2014), and it was found that disrupting the associated areas in primary motor cortex with repetitive TMS impairs categorical perception of artificial phonetic stimuli (Möttönen and Watkins 2009).

In sum, there are an increasing number of studies that demonstrate an essential function of motor system activation for auditory perception. This evidence goes against the dual-stream model, which only expects a modulatory role, and supports motor resonance theory and sensorimotor theory. So far research has focused on sound categories related to human action, and so it is difficult to arbitrate between these two latter theories. Future work could specifically investigate whether motor system activation is also necessary for the adequate functioning of auditory perception in the case of understanding non-action-related environmental sounds. Moreover, if sensorimotor theory is on the right track, we should expect that there are disruptions of the motor system that specifically impair the sensorimotor contingencies that constitute the "feel" of auditory experiences such they are no longer fully experienced as auditory. These interventions could be a target for future research.

4.3 The role of expertise in motor system activation

Most of the evidence for a role of expertise in motor system activation comes from studies of passive listening to music and rhythms. For example, an EEG study found evidence suggesting that there is increased motor system activity during passive listening to music that one can play and to music with similar sequences (Schalles and Pineda 2015). Similarly, motor excitability has been found to be increased for rehearsed compared to non-rehearsed musical pieces (D'Ausilio et al. 2006). An fMRI study comparing professional musicians with non-musicians in a task involving passive listening to short piano melodies revealed that the former showed increased activation in a distributed cortical network that also included supplementary motor and premotor areas (Bangert et al. 2006). Another fMRI study of non-musicians engaging in passive listening of music consistently found activation of premotor areas, especially when participants had been trained to play the music by ear (Lahav et al. 2007).

The effects of expertise can be rapid, as revealed by a couple of TMS studies which found that newly learnt action-related sounds (pure tones produced by button presses) activated the motor system in a goal-dependent manner, but this activation was muscle-independent and not somatotopically organized (Ticini et al. 2012, 2019). In addition to the rapid onset of the effects of expertise, the effects are persistent, as revealed by a TMS study of motor-evoked potentials following short audiomotor training sessions (Ticini et al. 2017). Several brain imaging studies comparing musicians and non-musicians found increased functional integration among perceptual and motor networks even during the resting state (Luo et al. 2012; Palomar-García et al. 2017).

There are also seems to be a more general effect of familiarity. An fMRI study with non-musicians found the supplementary motor cortex to be especially activated by familiar sounds, and also found the motor cortex to be more active for liked compared to disliked music (Silva Pereira et al. 2011). Such appraisal may also interact with expertise. A motor-evoked potential study found stronger engagement of the motor system during passive listening to high-groove music, particularly in musicians (Stupacher et al. 2013).

The effects of expertise are not limited to music and rhythms. An fMRI study revealed that when expert athletes are compared to novices they have greater activation of motor planning areas when passively listening to action-related sounds, especially if they were familiar sport-related sounds (Woods et al. 2014). Interestingly, novice athletes, on the other hand, showed greater activation for unfamiliar sounds, whether from sports or other actions on the environment. These results are somewhat unexpected from the perspective of motor resonance theory, given that greater activation of the motor system is expected for the perception of others' actions that are familiar and hence can more directly activate the perceiver's corresponding motor repertoire.

Sensorimotor theory has more diverse conceptual resources to deal with these kinds of results. On the one hand, in accordance with the principle of corporality, it expects that skilled perceivers have greater motor system activation, which reflects their broader and more fine-grained know-how of how the sounds of their expert activity tend to change with respect to bodily movement. Moreover, given that this motor system activation reflects the perceiver's own sensorimotor skill, rather than the neural activity underlying the perceived action of the other, it also expected that this skill generalizes to contexts not directly related with the expert activity to the extent that the contexts have overlapping sensorimotor contingencies. On the other hand, it can also account for increases in motor system activation in response to the perception of unfamiliar sounds, which reflects their increased alerting capacity, leading to a poised bodily state, especially for corresponding head- and gaze orientation. We also expect that novices are particularly susceptible to being alerted by unfamiliar sounds as they are actively trying to learn to distinguish relevant from irrelevant stimuli.

4.4 The scope of motor system activation

Motor system activation during perception of speech, music, and more generally action-related sounds is well established, and there is some evidence that the scope of activation is restricted to categories of sounds that are within the perceiver's motor repertoire. For example, an ERP-study found that during semantically primed auditory perception premotor areas are modulated by action-related sounds (like clapping hands), but not by non-action-related sounds (like boiling water) (Pizzamiglio et al. 2005). An fMRI study of passive listening to different sound categories found that non-vocal human action-related sounds led to preferential activation of motor-related networks, while animal sounds, which represented a category of non-vocal non-human action-related sounds that cannot be fully reproduced by humans, evoked a lesser degree of audio-motor associations (Engel et al. 2009); mechanical sounds, which were judged to be independent of a human instigating the action, and environmental sounds such as rain, wind, and fire preferentially activated cortical networks different from regions related to the motor system and mirror neuron system.

But there is also conflicting evidence for a broader scope of motor system activation. An fMRI study of long runs of broadband noise found that "nonauditory" areas, including areas well known for their role in motor planning and goal-directed behavior, showed significant modulations of activity at the rate of the sound stimulation cycles (Langers and Melcher 2011). They therefore conclude that the "results suggest that nonauditory centers play a role in sound processing at a very basic level, even when the sound is not intertwined with behaviors requiring the well-known functionality of these regions" (ibid.: 233). While they highlight that their findings do not conclusively demonstrate a functional role, they also consider that "it might be more reasonably argued that all observed sound-evoked responses have functional significance, in which case the nonauditory areas recruited by sound may influence how the sound itself is processed and/or consciously perceived" (ibid.: 242). A functional interpretation is further supported by the finding that under adverse listening conditions, such as would be induced by noise, increased activity in the speech motor system may compensate for the loss of specificity in the auditory system via sensorimotor integration (Du et al. 2014). Sensorimotor theory predicts that the involvement of these motor areas partially reflects the listener's skill of knowing how the sound sensations would change as a result of potential bodily movements.

This finding of motor system activation in response to listening to noise significantly goes beyond the explanatory scope of motor resonance theory. Moreover, even classical mirror neuron studies of auditory perception were already suggestive of a broader scope of motor system activation, although this evidence has received less attention in the literature. For example, a study involving intracranial recordings from three awake monkeys assessed the effect of non-action-related sounds, namely arousing and emotional sounds such as loud noises and other animal calls, on 32 neurons from the ventral premotor cortex (area F5), and found a response, albeit somewhat reduced in magnitude: five of those neurons were weakly active, while the rest did not respond significantly (Kohler et al. 2002). Still, a reduced response is not an absence of response, a point to which we will return later.

In the case of humans, it is useful to consider in detail an EEG study of non-musicians' passive listening to action-related sounds of the non-natural environment, which found stronger evidence for motor system activation for sounds that typically cue a further responsive action on the part of the listener, even though all sounds were familiar, and all actions were supposedly within the participants' motor repertoire (De Lucia et al. 2009). This finding sits uncomfortably with strict definitions of the mirror mechanism, and accordingly the authors interpret their finding as consistent with the "broadly congruent" class of mirror neurons. They also consider that the stronger motor system activation could reflect listeners' ability to better understand the intentions of others based on a recognition of the situation in which the action-cuing sounds were produced, but this interpretation does not seem plausible given that several of the action-cuing sounds do not involve clear action goals.

More specifically, the study involved 10 sounds of what the authors deemed "context-free" actions, namely exemplars of notes being played on the following musical instruments: accordion, flute, guitar, harmonica, harp, organ, piano, saxophone, trumpet, and violin (i.e. both string and brass instruments involving mouth and hand actions). The authors emphasize that these stimuli were neither rhythmic nor melodic in character and were not perceived as music, but rather in terms of the

instrument generating the sound. These sounds were contrasted with 10 sounds of "context-related" actions included exemplars of the following objects: bicycle bell, car horn, cash register, cuckoo clock, doorbell, closing door, glass shattering, police siren, church bell, and telephone (i.e. sounds that typically trigger a responsive action upon being heard).

We note that both a cuckoo clock and a church bell are typically triggered by an automatic mechanism, while other examples are only indirectly caused by another person's action (i.e. the ongoing sound of a police siren or telephone ringing were originally triggered by a person's action, but their continued sounding does not depend on further actions), and two also have a non-intentional natural causes (e.g. winds often cause a door to close, and can also cause a glass to fall and shatter). In other words, at least half of the context-related sounds have no clear causal relation to another's concurrent movements and consequently neither to their action goals. Our interpretation of these acoustic stimuli is consistent with other studies that have used similar sounds in a category of sounds that cannot be generated by human actions (Galati et al. 2008). According to motor resonance theory we should therefore expect absent or at least significantly decreased motor system involvement for this category of environmental sounds, compared to a category of sounds generated by actions on instruments, and yet we find exactly the opposite. We therefore interpret this study as demonstrating motor system activation also for the perception of sounds that were not caused by actions, as expected by sensorimotor theory but not by the competing theories.

5 Discussion

All theories of auditory perception agree that neural activity is required for a human or animal to be able to experience the sounds of their environment. However, while more established theories have already developed detailed models of this neural activity, the ramifications of sensorimotor theory for neuroscience are still largely to be explored. Here we make a tentative proposal about the neural basis of the exercise of audio-motor know-how, and compare it to the predictions of the other theories, which leads to a unique prediction regarding the function and scope of motor system activation. But this is not the full story, as sensorimotor theory claims that perceiving is something done by the whole organism, and therefore cannot be reduced to brain activity. We therefore briefly consider different lines of research into auditory physiology, which are converging on a theory of audition as an inherently active process involving forms of movement even at the level of the peripheral auditory system.

5.1 Dual stream model, motor resonance theory, or sensorimotor theory?

Not all passive listening studies report motor system activation, and this inconsistency in the literature fits with the dual stream model's claim that this activation is only secondary rather than essential for auditory perception. However, at this point it is still not clear whether this inconsistency is best interpreted as absence of evidence or evidence of absence. As others have highlighted, motor system involvement in speech perception has not frequently been sought until recently (Galantucci et al. 2006), and it is still commonly overlooked in the context of auditory perception more generally

(Lima et al. 2016). Moreover, not all studies of passive listening are designed to detect motor system involvement. For example, an fRMI study of a trance state of consciousness induced by rhythmic drumming did not report any evidence of motor system activation, but this lack of evidence becomes understandable when we consider that the authors were only interested in comparing trance versus non-trance conditions that both involved auditory perception (Hove et al. 2016). More importantly, accumulating evidence is suggestive of essential motor system involvement in auditory perception, especially for speech, music, rhythm, and other action-related sounds, and this involvement is often somatotopically organized. This evidence favors motor resonance theory over the dual stream model.

The relationship between motor resonance theory and sensorimotor theory is more complex. It could be argued that both theories target different aspects of auditory experience, with sensorimotor theory focusing on the specifically auditory quality of the perceptual experience, while motor resonance theory focuses on the understanding of action goals. Accordingly, both theories could simply coexist. However, sensorimotor theory's rejection of appeals to neural representations, and motor resonance theory's appeal to motor representations, suggest that this coexistence would not be coherent. Accordingly, it seems that sensorimotor theory would be better off trying to account for the existing evidence on its own terms, even though this is not a straightforward task. Nevertheless, it should also be noted that, from the perspective of sensorimotor theory, there are several issues with the typical designs of mirror neuron studies that imply that the existing evidence should not always be taken at face value:

(1) *The preferred choice of non-action-related sounds as a control condition could hide broader motor system involvement.* This choice of a control condition that involves auditory perception makes sense from the perspective of motor resonance theory, given that it expects motor system activation for action-related sounds only. From the perspective of sensorimotor theory, however, the only fully appropriate control condition would be a complete absence of auditory events. Under sensorimotor theory's proposal of a general involvement of motor processes in auditory perception as such, this creates a special methodological difficulty for using fMRI to look for motor system activation during passive listening, because it is difficult to ensure that participants do not inadvertently experience sounds produced by the scanner's operation during control conditions.

(2) *Increased motor system activation for action-related sounds does not entail the absence of motor system involvement for non-action-related sounds.* Typically, results are interpreted in favor of motor resonance theory if there is significantly more activation of the motor system for action-related compared to non-action-related sounds. However, this means that evidential support for the restricted scope of motor resonance theory over the broad scope of sensorimotor theory is mainly a matter of degree. This ambiguous situation is consistent with the previous point, namely that non-action-related sounds are questionable choices for control conditions if they are supposed to represent conditions with absence of motor system activation. Differences in degree of activation may therefore be better accounted for in terms of differences in expertise or familiarity, or in terms of differences in alerting capacity and corporality.

(3)   *The focus of analysis has been mostly restricted to motor areas somatotopically related to the actions that produced the perceived sounds.* This restriction could hide a broader involvement of the motor system. Moreover, this restriction only makes sense in the context of studies of action-related sounds, which makes it difficult to assess the relative importance of these somatotopic areas compared to other motors areas in other contexts. In particular, from the perspective of sensorimotor theory, we do not expect motor system activation to always be constrained by somatotopic congruency: perceiving is about exercising know-how of audio-motor contingencies that depend on the self's potential movements rather than on the other's actual actions.

(4)   *The criterion of congruency is not applied consistently.* It has long been known that some neurons of the motor system in monkeys activate even though they do not correlate with any of the animal's movements, while others are only "logically related" to the perceived action because of their involvement in other kinds of actions (di Pellegrino et al. 1992). Gallese et al. (1996: 600–601) even partition "mirror neurons into three broad classes: 'strictly congruent', 'broadly congruent' and 'non-congruent'" with the perceived action. However, rather than accepting neurons with only logically related or even non-congruent activity as mirror neurons, this evidence points to broader motor system involvement.

Accordingly, there are important methodological and theoretical biases in the mirror neuron literature that stack the odds against sensorimotor theory. At the same time, given that the current empirical record does not consistently favor motor resonance theory even despite these biases, sensorimotor theory may nevertheless still emerge as a robust alternative theory of auditory perception. We highlight three points of contention:

a).   *The role of somatotopic organization:* Some, but not all, studies find motor system activation to be somatotopically organized. For standard sensorimotor theory, this somatotopic organization is not expected, but it can explain its occurrence as resulting from the perceiver's history of audio-motor associations (Cook et al. 2014), especially if the actions are in the perceiver's motor repertoire. These activations may also play a functional role in poising the perceiver to respond to what is perceived. Indeed, there are efforts to extend sensorimotor theory into the social domain in terms of self-other contingencies: others' actions are perceived as affordances for interaction, which shape how the other's actions are experienced (McGann and De Jaegher 2009). This response can include more abstract actions, because affordances are not limited to sensorimotor contingencies, but include the perceiver's whole range of action capabilities (Rietveld et al. 2018). However, this would still leave it unexplained why this poised response tends to preferentially activate parts of the nervous system specifically related to the effector of the perceived action. One possibility would be to consider in more detail the affordances of imitative or matching responses, e.g. speech tends to be responded to with more speech, and it is likely that when one perceives another person eating that one will have the affordance of eating, too. Another possibility that could be further developed in future work is to appeal to recent work in the enactive approach, which proposes that perception and imagination are deeply unified (Bruineberg et al. 2019; Kirchhoff 2018). Even basic sensory imaginings are best

understood as re-creative reenactments enabling appropriate interactions (Hutto 2015). The aim would therefore be to develop a theory of the neural underpinnings of empathy and shared affectivity in terms of the perceiver's responsiveness to social affordances (Kiverstein 2015). One compelling possibility is that neural activity is coordinated across perceiver and perceived because they tend to become integrated into an overarching sensorimotor interaction process (Froese 2018) that is organized and ordered by interpersonal synergies (Chemero 2016).

b). *The scope of motor system activation*: Some, but not all, studies find a lack of motor system activation for non-action-related sounds. The study by De Lucia et al. (2009) is of particular interest as a clear exception because at least half of their action-cuing sounds have no clear causal relation to others' movements, and yet these stimuli are associated with increased motor system activation compared to sounds of instrumental actions. Sensorimotor theory provides a more coherent account of these results. First, it can easily account for the fact that both categories evoked some motor system activation, because all auditory events are characterized by corporality. In other words, their phenomenological quality as specifically auditory events depends on the auditory modality's characteristic sensorimotor contingencies. Second, and more importantly, it can straightforwardly account for the stronger motor system activation of the action-cuing sounds in terms of the alerting capacity or grabbiness of the action-cuing auditory events. Many of the sounds of this category are specifically designed to be particularly good at alerting listeners and grabbing their attention: horns, bells, sirens, alarms, etc.

c). *The role of perceptual modality:* Some, but not all, studies find modality-independent motor system activation. For instance, early work with monkeys highlighted that 11 out of 22 audiovisual mirror neurons did not behave differently when comparing unimodal (vision or audition) and multimodal (audiovisual) stimuli, which was taken to be in support of the mirror neuron mechanism; but already in that study a second category of 8 neurons was characterized by the fact that the strongest response was observed when the sound and vision of the action were presented together (Keysers et al. 2003). Moreover, a TMS study of passive listening to action-related-sounds, which was inspired by this second class of neurons, found a selective response increase in primary motor cortex excitability for audiovisual stimuli compared to unimodal stimuli, albeit only for cross-modally coherent multimodal stimuli (Alaerts et al. 2009). The authors interpret their results as favoring a modality-dependent activation of the motor system, which is challenging for motor resonance theory's expectation that activation should be modality independent. This kind of finding is also challenging for theories based on sensory prediction, such as the dual stream model, because the redundancy provided by multiple modalities should ease processing and hence lower motor system activity (Gordon et al. 2018b). We propose that a sensorimotor theory can account for the diversity of findings: modality-independent activation suggests that these neurons are involved in sensorimotor contingencies that are structurally similar across modalities, whereas modality-dependent increases in activation suggests that the neurons are involved in sensorimotor contingencies that are specific to each of the perceptual modalities involved and hence must be realized in parallel. In addition, a violation of expectation of cross-modal congruence is expected to heighten the alerting capacity of the stimuli.

Accordingly, many classical empirical findings of motor resonance theory can be incorporated into sensorimotor theory, albeit stripped of their other's-action-goals-targeted motor representational interpretation. Future work could expand the list of topics we discussed above. For example, sensorimotor theory is starting to have the conceptual resources to even challenge motor resonance theory in its core domain of social cognition. Evidence that empathy correlates with motor system activation (e.g. Gazzola et al. 2006) could be re-interpreted in terms of extensions of sensorimotor theory into the social domain: we expect that there will be motor system involvement in the perceiver's exercise of mastery of "self-other contingencies" (McGann and De Jaegher 2009) and of "sensorimotor empathy" (Chemero 2016). This would enable us to give a nonrepresentational account of the evidence of the mirror mechanism for action understanding in terms of the neurodynamical basis of the embodied skills of social coordination (Di Paolo and De Jaegher 2012) and direct social perception (Gallagher 2007).

5.2 Neural basis of the exercise of audio-motor know-how?

According to sensorimotor theory, perceivers tend to be poised to confirm their acquired familiarity with how sensations would change with respect to potential movements, a state of readiness sometimes referred to as the "attunement" to those sensorimotor laws (Foglia and O'Regan 2016: 192). It follows that compatibility with this key tenet of sensorimotor theory requires that the neural processes enabling the perceiver's attunement must be characterized by a relatively high degree to which these processes are influenced by potential interactions with the environment (Degenaar and O'Regan 2015). This requirement suggests an alternative interpretation of the discovery that a large number of cortical areas involved in action execution are also active during action observation: rather than construing this basic principle of brain functioning as a "mirror mechanism" (Rizzolatti and Sinigaglia 2016), it can instead be interpreted as an important part of the sensorimotor processes that constitute perceptual access to the environment, including to the social domain in terms of self-other contingencies and perceptual access to others' action goals. Our review of the literature on passive listening tasks permits us to refine this general proposal of the role of motor system involvement.

We suggest that a perceiver's exercise of the mastery of audio-motor contingencies is partially enabled by complex patterns of neural activity that tend to be realized in the supplementary motor areas (SMAs), which consist of two distinct fields: one anterior, the pre-supplementary motor area (pre-SMA), and one posterior, the supplementary motor area proper (SMA).[4] Although these regions are more typically investigated in relation to their motor functions, they are also known for other functions, and in particular they are consistently reported in studies of auditory perception and auditory imagery even when the tasks do not involve overt movement. This

---

[4] Note that we are not committed to claiming that these particular areas per se are necessary for the exercise of sensorimotor mastery. What is required are specific kinds of neural activity patterns that emerge in response to sensory stimuli (Luczak et al. 2009), and presumably these complex patterns tend to be more easily realized with the help of these areas because they acquired a particularly suitable neural connectivity.

involvement is commonly overlooked, but Lima et al. (2016) review converging evidence that lead them to claim that the SMAs play a role in (1) facilitating spontaneous motor responses to sound, and in (2) supporting a flexible engagement of sensorimotor processes to guide auditory experience and to enable auditory imagery. We propose that these two hypothesized roles make a particularly fitting match with sensorimotor theory's characterization of neural activity associated with the exercise of sensorimotor skill in terms of two special properties (O'Regan et al. 2005), namely its (1) "alerting capacity" (also known as "grabbiness"), and (2) its "corporality" (or "bodiliness"), respectively.

A systematic study of the effects of direct electrical stimulation of the SMAs (Fried et al. 1991) are consistent with this interpretation. Stimulation tended to coactivate larger and more distributed groups of muscles than in the primary motor area, which is suggestive of a more holistic role, as would be expected for the basis of a bodily skill. In addition, the responses to stimulation were also frequently situated more at the abstract level of movement intentions rather than of concrete movement execution. People often reported an "urge" to move a part of their body or an anticipation that such a movement was about to occur, but without necessarily following through with the overt movement. In addition, some people reported subjective experiences of bodily movement even though no overt movement was observed by the experimenters. Taken together, this range of effects elicited by direct stimulation of the SMAs fit with sensorimotor theory's claim that perceiving is a bodily skill involving the exercise of sensorimotor mastery which, importantly, after having been acquired normally only involves *potential intentional* movements rather than actual intentional movements, and yet can still partially constitute subjective experience.

Here intentional movement is to be contrasted with accidental or goalless movement; it does not imply explicit motor planning or rehearsal. In fact, the SMAs are activated during passive listening even in the absence of explicit anticipation (Chen et al. 2008). On the other hand, clinical research has shown that implicit anticipatory postural control, including neck flexion and neck posture, is particularly associated with the SMAs and premotor areas (Fujiwara et al. 2009; Naumann et al. 2000). The SMA's association with the neck muscles is particularly relevant, because potential neck movements are arguably the most important kinds of movements both for the alerting capacity and corporality of auditory perception. In addition, the alerting capacity is also likely to recruit gaze shifts, which have been associated with an area located right next to the SMAs, namely Brodmann Area 8 (Lanzilotto et al. 2013).

Note that this proposal differs crucially from the functional role that the dual stream model and motor resonance theory assign to the activation of the SMAs during motionless listening. They both interpret its activation in terms of the perceiver's motor representation of an action that produced the perceived sound. They differ in how they envision the function of that motor representation, i.e. as either in the service of predicting which sounds will follow (dual stream model) or of directly matching the action goal (motor resonance theory). In contrast, sensorimotor theory rejects this representational interpretation and instead proposes that what is at stake in this motor activation is not a motor representation of a sound-producing action, but rather the perceiver's own activity, namely the perceiver's skillful sensorimotor engagement with

the acoustic environment.[5] One immediate implication of this interpretation is that for sensorimotor theory the scope of functional activation of the SMAs is not limited to perceiving speech- and action-related sounds, but is involved in auditory experience as such. We therefore suggest that research into the neural basis of passive listening should broaden their current focus on speech, music, and rhythms to include a wider range of acoustic stimuli, especially sounds that are typically not under gestural control (e.g. environmental sounds like hearing a plane fly overhead).

5.3 Auditory perception as a bodily activity?

Like other theories of perception, sensorimotor theory accepts that neural activity forms an important basis of perceptual experience, but in contrast to most other theories it does not limit itself to neural activity: it extends its focus outward to the dynamic relation between brain, body, and world (Hurley and Noë 2003; Di Paolo et al. 2017). It claims that perception is a bodily doing or bodily activity, a term which refers to the specific embodiment of the perceiver and their capacity for both covert and overt movements (Myin 2016). It thereby acknowledges that the sensorimotor contingencies characteristic of each modality are shaped by the morphology and physiology of the organs that are involved, and hence that a complete account of auditory experience must also take into account the structural effects of auditory physiology (Hofman et al. 1998), as well as the complex activity of the peripheral auditory system (Manley et al. 2008). In fact, even the function of the rest of the body could be considered, for example the possible role of facial muscle tension in auditory attention (Cohen et al. 1992), and more generally the complex and plastic feedback circuitry between the motor cortex and the spinal cord (Graziano 2006). A complete assessment of all of this additional evidence goes beyond this review, which was focused on the role of the central nervous system. Nevertheless, we mention some tantalizing possibilities for future research in this direction.

- *Ear motility:* Movement of the outer ears plays an important role in audition for many animals, but human outer ears have become largely immobile. Nevertheless, even human audition involves plenty of motility in other parts of the ears. For example, the hair bundle is a mechanosensitive organelle that protrudes from the hair cell and enables it to respond to auditory frequencies. But this organelle can also convert energy of biochemical origin into mechanical work and thus act as a small engine, and this motility has implications for the detection of minute oscillatory stimuli (Martin 2008). In addition, there are several motor mechanisms that modify the auditory transduction process in the peripheral hearing system (Gruters et al. 2018). For example, skeletal muscles attached to the ossicles contract in response to loud sounds to attenuate vibrations, but they are also involved in enhancing sound discrimination (Borg and Counter 1989). Contractions in response to novel acoustic stimuli and contractions associated with other motor activity, such as chewing or

---

[5] To be fair, it is possible to model this sensorimotor engagement in counterfactual terms, namely in terms of internal models that predict the sensory consequences of potential movements (Seth 2014). However, this kind of approach sits uneasily with sensorimotor theory's avoidance of appealing to internal representations (Flament-Fultot 2016; Froese 2014; Silverman 2018).

vocalization, provide a means to separate auditory stimuli from their background and a mechanism facilitating stable attention to a continuous sound (Simmons 1964).

- *Otoacoustic emissions:* Sounds are generated in the ear canal as a result of normal physiological activity in the cochlea, and these are therefore used in diagnosing healthy middle ear functioning. There is a mechanism that converts electrical signals into mechanical oscillation of the cochlea's outer hair cells (electromotility). These sounds have been related to reduced damping and improved sensitivity. One possibility is active amplification used by the auditory system to detect sounds that carry less energy than the background (Camalet et al. 2000). They could also be essential for maintaining stable performance in the auditory system in terms of self-regulation of the amount of positive auditory feedback (Kemp 2008). Activation of the efferent auditory system, which is argued to be responsible for otoacoustic emissions (Guinan 2010), has also been correlated with improved sound-localization performance under noise conditions (Andéol et al. 2011). The eardrum also spontaneously oscillates in a way that produces sounds with an amplitude and phase dependent on the direction and horizontal amplitude of eye saccades, possibly creating eye-movement-related binaural cues for auditory-visual integration (Gruters et al. 2018).

- *Auditory-vestibular integration*: The vestibular system is involved in perceiving head and body orientation in space (Gdowski and McCrea 1999), in sound localization (Wallach 1940), and is differentially responsive during active versus passive head movements (Cullen and Roy 2004). The auditory and vestibular systems are anatomically joined, providing the basis for functional auditory-vestibular integration. For instance, sound pressure invariably activates both auditory and vestibular systems, thereby facilitating multisensory integration for the spatial localization of sound (Oh et al. 2018). This integration can be used for vestibular spatial recalibration by training participants to rely on their perception of a distal sound source as a cue for their spatial orientation (Schumann and O'Regan 2017). In addition, vestibular activity may play a key role in the effect of head movement on how rhythms are experienced (Phillips-Silver and Trainor 2008).

In general, the active processes of the ears and their biological basis in different forms of motility and sound generation have become an important area of auditory research (Manley et al. 2008; Maoiléidigh and Ricci 2019). This area is a natural complement for sensorimotor theory's claim that perceiving is a skillful activity of the whole organism, which implies that, like in the case of other bodily doings, the act of perceiving is not only based on brain activity, but also on bodily activity. Indeed, there is some initial evidence that even such bodily activity in the peripheral auditory system can be subject to skill, like musicianship (Bidelman et al. 2016; Bidelman et al. 2017). Future work could investigate in more detail the roles that these active processes of the peripheral auditory system play in the perceiver's exercise of their mastery of audio-motor contingencies.

6 Conclusion

In the end, from the perspective of sensorimotor theory, there is little passivity about "passive listening" tasks. First of all, as Myin (2016) has argued at length, perceiving is

something that is actively done by an organism, often via overt bodily movement, although this activity is not limited to such movement: "Holding still is something we do, but obviously without moving – if we move, we stop or fail to hold still" (p. 88). Indeed, participants have a hard time complying with the instruction to remain motionless. Second, sensorimotor theory leads to the hypothesis that the development of the auditory modality involves mastering the relevant audio-motor contingencies via active bodily engagement with the environment during ontogenetic development, and this history of sensorimotor engagement is constitutive of the feel of normal auditory experience, as reflected by consistent motor system activation even during passive listening tasks. Third, even the experience of simply being acoustically stimulated while being motionless is enabled by complex bodily activity: "a specific and active doing in which the body reacts to a specific kind of stimulation" (p. 90), as revealed by the fact that even inner ear activity behaves in a contextualized and sensitive manner.

We conclude that defending the core claim of sensorimotor theory, namely that perceiving is a kind of bodily doing, ultimately requires that future work on auditory experience takes more seriously the constitutive role played by physiological activity itself, of which neural activity is an important part. Indeed, sensorimotor theory provides a compelling account of why there is consistent motor system involvement during listening even without overt movement. We have proposed that the pre-SMA's key auditory roles, namely in (1) facilitating spontaneous motor responses to sound and in (2) supporting a flexible engagement of sensorimotor processes to guide auditory experience and enable auditory imagery, make a particularly fitting match with sensorimotor theory's key concepts, namely its (1) "alerting capacity" (or "grabbiness"), and (2) its "corporality" (or "bodiliness"). But there is more to the body than just the brain: we have highlighted that there is an opportunity to develop sensorimotor theory into new directions in terms of the still relatively poorly understood active processes of the peripheral auditory system. In general, it is clear that sensorimotor theory holds promise for the development of a theory of auditory experience that can better integrate the full range of available evidence.

References

Aglioti, S. M., & Pazzaglia, M. (2010). Representing actions through their sound. *Experimental Brain Research, 206*, 141–151.

Aglioti, S. M., & Pazzaglia, M. (2011). Sounds and scents in (social) action. *Trends in Cognitive Sciences, 15*(2), 47–55.

Aizawa, K. (2019). Is perceiving bodily action? *Phenomenology and the Cognitive Sciences*. https://doi.org/10.1007/s11097-018-9592-9.

Alaerts, K., Swinnen, S. P., & Wenderoth, N. (2009). Interaction of sound and sight during action perception: evidence for shared modality-dependent action representations. *Neuropsychologia, 47*(12), 2593–2599.

Andéol, G., Guillaume, A., Micheyl, C., Savel, S., Pellieux, L., & Moulin, A. (2011). Auditory efferents facilitate sound localization in noise in humans. *The Journal of Neuroscience, 31*(18), 6759–6763.

Aytekin, M., Moss, C. F., & Simon, J. Z. (2008). A sensorimotor approach to sound localization. *Neural Computation, 20*, 603–635.

Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinze, H. J., & Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage, 30*(3), 917–926.

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–645.

Beaton, M. (2013). Phenomenology and embodied action. *Constructivist Foundations, 8*(3), 298–313.

Beaton, M. (2016). Sensorimotor direct realism: how we enact our world. *Constructivist Foundations, 11*(2), 265–297.

Bidelman, G. M., Nelms, C., & Bhagat, S. P. (2016). Musical experience sharpens human cochlear tuning. *Hearing Research, 335*, 40–46.

Bidelman, G. M., Schneider, A. D., Heitzmann, V. R., & Bhagat, S. P. (2017). Musicianship enhances ipsilateral and contralateral efferent gain control to the cochlea. *Hearing Research, 344*, 275–283.

Bishop, J. M., & Martin, A. O. (2014). Contemporary sensorimotor theory: a brief introduction. In J. M. Bishop & A. O. Martin (Eds.), *Contemporary sensorimotor theory* (pp. 1–22). Switzerland: Springer.

Borg, E., & Counter, S. A. (1989). The middle-ear muscles. *Scientific American, 261*(2), 74–81.

Bornkessel-Schlesewsky, I., Schlesewsky, M., Small, S. L., & Rauschecker, J. P. (2015). Neurobiological roots of language in primate audition: common computational properties. *Trends in Cognitive Sciences, 19*(3), 142–150.

Brefczynski-Lewis, J. A., & Lewis, J. W. (2017). Auditory object perception: a neurobiological model and prospective review. *Neuropsychologia, 105*, 223–242.

Bruineberg, J., Chemero, A., & Rietveld, E. (2019). General ecological information supports engagement with affordances for 'higher' cognition. *Synthese.* https://doi.org/10.1007/s11229-018-1716-9.

Buccino, G., Riggio, L., Melli, G., Binkosfki, F., Gallese, V., & Rizzolatti, G. (2005). Listening to action-related sentences modualtes the activity of the motor system: a combined TMS and behavioral study. *Cognitive Brain Research, 24*, 355–363.

Camalet, S., Duke, T., Jülicher, F., & Prost, J. (2000). Auditory sensitivity provided by self-tuned critical oscillations of hair cells. *Proceedings of the National Academy of Sciences of the USA, 97*(7), 3183–3188.

Caramazza, A., Anzellotti, S., Strnad, L., & Lingnau, A. (2014). Embodied cognition and mirror neurons: a critical assessment. *Annual Review of Neuroscience, 37*, 1–15.

Chabris, C. F., & Simons, D. J. (2010). *The invisible Gorilla: and other ways our intuition deceives us*. London: Harper Collins.

Chemero, A. (2016). Sensorimotor empathy. *Journal of Consciousness Studies, 23*(5–6), 138–152.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*(12), 2844–2854.

Cohen, B. H., Davidson, R. J., Senulis, J. A., Saron, C. D., & Weisman, D. R. (1992). Muscle tension patterns during auditory attention. *Biological Psychology, 33*(2–3), 133–156.

Cook, R., Bird, G., Catmur, C., & Press, C. (2014). Mirror neurons: from origin to function. *Behavioral and Brain Sciences, 37*(2), 177–192.

Cooke, E., & Myin, E. (2011). Is trilled smell possible? How the structure of olfaction determines the phenomenology of smell. *Journal of Consciousness Studies, 18*(11–12), 59–95.

Cullen, K. E., & Roy, J. E. (2004). Signal processing in the vestibular system during active versus passive head movements. *Journal of Neurophysiology, 91*(5), 1919–1933.

D'Ausilio, A., Altenmüller, E., Olivetti Belardinelli, M., & Lotze, M. (2006). Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. *European Journal of Neuroscience, 24*(3), 955–958.

D'Ausilio, A., Pulvermüller, F., Salmas, P., Bufalari, I., Begliomini, C., & Fadiga, L. (2009). The motor somatotopy of speech perception. *Current Biology, 19*, 381–385.

D'Ausilio, A., Maffongelli, L., Bartoli, E., Campanella, M., Ferrari, E., Berry, J., & Fadiga, L. (2014). Listening to speech recruits specific tongue motor synergies as revealed by transcranial magnetic stimulation and tissue-Doppler ultrasound imaging. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 369*(1644), 20130418. https://doi.org/10.1098/rstb.2013.0418.

Dalton, P., & Fraenkel, N. (2012). Gorillas we have missed: sustained inattentional deafness for dynamic events. *Cognition, 124*, 367–372.

Dannenbring, G. L. (1976). Perceived auditory continuity with alternately rising and falling frequency transitions. *Canadian Journal of Psychology, 30*(2), 99–114.

De Lucia, M., Camen, C., Clarke, S., & Murray, M. M. (2009). The role of actions in auditory object discrimination. *Neuroimage, 48*(2), 475–485.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., Mazziotta, J. C., & Fazio, F. (1994). Mapping motor representations with positron emission tomography. *Nature, 371*, 600–602.

Degenaar, J., & O'Regan, J. K. (2015). Sensorimotor theory of consciousness. *Scholarpedia, 10*(5), 4952.

Degenaar, J., & O'Regan, J. K. (2017). Sensorimotor theory and enactivism. *Topoi, 36*, 393–407. https://doi.org/10.1007/s11245-015-9338-z.

Di Paolo, E. A., & De Jaegher, H. (2012). The interactive brain hypothesis. *Frontiers in Human Neuroscience, 6*(163). https://doi.org/10.3389/fnhum.2012.00163.

Di Paolo, E. A., Buhrmann, T., & Barandiaran, X. (2017). *Sensorimotor life: An enactive proposal*. Oxford: Oxford University Press.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research, 91*, 176–180.

Dreyer, F. R., Frey, D., Arana, S., von Saldern, S., Picht, T., Vajkoczy, P., & Pulvermüller, F. (2015). Is the motor system necessary for processing action and abstract emotion words? Evidence from focal brain lesions. *Frontiers in Psychology, 6*(1661). https://doi.org/10.3389/fpsyg.2015.01661.

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings of the National Academy of Sciences of the USA, 111*(19), 7126–7131.

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2016). Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nature Communications, 7*(12241). https://doi.org/10.1038/ncomms12241.

Engel, L. R., Frum, C., Puce, A., Walker, N. A., & Lewis, J. W. (2009). Different categories of living and non-living sound-sources activate distinct cortical networks. *Neuroimage, 47*, 1778–1791.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excibtability of tongue muscles: a TMS study. *European Journal of Neuroscience, 15*(2), 399–402.

Flament-Fultot, M. (2016). Counterfactuals versus constraints: towards an implementation theory of sensori-motor mastery. *Journal of Consciousness Studies, 23*(5–6), 153–176.

Foglia, L., & O'Regan, J. K. (2016). A new imagery debate: Enactive and sensorimotor accounts. *Review of Philosophy and Psychology, 7*, 181–196.

Fowler, C. A., & Xie, X. (2016). Involvement of the speech motor system in speech perception. In P. van Lieshout, B. Maassen, & H. Terband (Eds.), *Speech motor control in Normal and disordered speech: Future developments in theory and methodology* (pp. 1–24). Rockville: ASHA Press.

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., Spencer, S. S., & Spencer, D. D. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *The Journal of Neuroscience, 11*(11), 3656–3666.

Froese, T. (2014). Steps toward an enactive account of synesthesia. *Cognitive Neuroscience, 5*(2), 126–127.

Froese, T. (2018). Searching for the conditions of genuine intersubjectivity: From agent-based models to perceptual crossing experiments. In A. Newen, L. De Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition* (pp. 163–186). Oxford: Oxford University Press.

Froese, T., & Ziemke, T. (2009). Enactive artificial intelligence: investigating the systemic organization of life and mind. *Artificial Intelligence, 173*(3–4), 366–500.

Fuchs, T. (2018). *Ecology of the brain: The phenomenology and biology of the embodied mind*. Oxford: Oxford University Press.

Fujiwara, K., Tomita, H., Maeda, K., & Kunita, K. (2009). Effects of neck flexion on contingent negative variation and anticipatory postural control during arm movement while standing. *Journal of Electromyography and Kinesiology, 19*, 113–121.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review, 13*(3), 361–377.

Galati, G., Committeri, G., Spitoni, G., Aprile, T., Di Russo, F., Pitzalis, S., & Pizzamiglio, L. (2008). A selective representation of the meaning of actions in the auditory mirror system. *Neuroimage, 40*, 1274–1286.

Gallagher, S. (2007). Simulation trouble. *Social Neuroscience, 2*(3), 353–365.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2*(12), 493–501.

Gallese, V., & Sinigaglia, C. (2018). Embodied resonance. In A. Newen, L. De Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition* (pp. 417–432). Oxford: Oxford University Press.

Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain, 119*(2), 593–609.

Gazzola, V., Aziz-Zadeh, L., & Keysers, C. (2006). Empathy and the somatotopic auditory mirror system in humans. *Current Biology, 16*, 1824–1829.

Gdowski, G. T., & McCrea, R. A. (1999). Integration of vestibular and head movement signals in the vestibular nuclei during whole-body rotation. *Journal of Neurophysiology, 82*(1), 436–449.

Gilchrist, I. D., Brown, V., & Findlay, J. M. (1997). Saccades without eye movements. *Nature, 390*, 130–131.

Gordon, C. L., Cobb, P. R., & Balasubramaniam, R. (2018a). Recruitment of the motor system during music listening: An ALE meta-analysis of fMRI data. *PLoS One, 13*(11), e0207213. https://doi.org/10.1371/journal.pone.0207213.

Gordon, C. L., Iacoboni, M., & Balasubramaniam, R. (2018b). Multimodal music perception engages motor prediction: a TMS study. *Frontiers in Neuroscience, 12*(736). https://doi.org/10.3389/fnins.2018.00736.

Graziano, M. (2006). The organization of behavioral repertoire in motor cortex. *Annual Review of Neuroscience, 29*, 105–134.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences, 27*(3), 377–396.

Gruters, K. G., Murphy, D. L. K., Jenson, C. D., Smith, D. W., Shera, C. A., & Groh, J. M. (2018). The eardrums move when the eyes move: a multisensory effect on the mechanics of hearing. *Proceedings of the National Academy of Sciences of the USA, 115*(6), E1309–E1318.

Guinan, J. J. (2010). Cochlear efferent innervation and function. *Current Opinion in Otolaryngology & Head and Neck Surgery, 18*(5), 447–453.

Hauk, O., Shtyrov, Y., & Pulvermüller, F. (2006). The sound of actions as reflected by mismatch negativity: rapid activation of cortical sensory-motor networks by sounds associated with finger and tongue movements. *European Journal of Neuroscience, 23*(3), 811–821.

Heidegger, M. ([1927] 1996). *Being and Time* (J. Stambaugh, Trans.). Albany: State University of New York Press.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition, 92*, 67–99.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience, 8*(5), 393–402.

Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron, 69*(3), 407–422.

Hofman, P. M., Van Riswick, J. G. A., & van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature Neuroscience, 1*(5), 417–421.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): a framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849–937.

Hove, M. J., Stelzer, J., Nierhaus, T., Thiel, S. D., Gundlach, C., Margulies, D. S., van Dijk, K. R. A., Turner, R., Keller, P. E., & Merker, B. (2016). Brain network reconfiguration and perceptual decoupling during an absorptive state of consciousness. *Cerebral Cortex, 26*(7), 3116–31124.

Hurley, S., & Noë, A. (2003). Neural plasticity and consciousness. *Biology and Philosophy, 18*, 131–168.

Husserl, E. (1966). *Zur Phänomenologie des inneren Zeitbewusstseins (1893–1917)*. Den Haag: Martinus Nijhoff.

Hutto, D. D. (2015). Overly enactive imagination? Radically re-imagining imagining. *The Southern Journal of Philosophy, 53*(Spindel Supplement), 68–89.

Hutto, D. D., & Myin, E. (2013). *Radicalizing Enactivism: Basic minds without content*. Cambridge: The MIT Press.

Isel, F. (2001). How do we account for the absence of "change deafness"? *Behavioral and Brain Sciences, 24*(5), 988.

Kemp, D. T. (2008). Otoacoustic emissions: Concepts and origins. In G. A. Manley, R. R. Fay, & A. N. Popper (Eds.), *Active processes and Otoacoustic emissions in hearing* (pp. 1–38). New York: Springer.

Keysers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., & Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Experimental Brain Research, 153*(4), 628–636.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: an account of the mirror neuron system. *Cognitive Processing, 8*(3), 159–166.

Kirchhoff, M. (2018). Predictive processing, perceiving and imagining: is to perceive to imagine, or something close to it? *Philosophical Studies, 175*, 751–767.

Kirchhoff, M. D., & Froese, T. (2017). Where there is life there is mind: in support of a strong life-mind continuity thesis. *Entropy, 19*(4), 169. https://doi.org/10.3390/e1904016.

Kiverstein, J. (2015). Empathy and the responsiveness to social affordances. *Consciousness and Cognition, 36*, 532–542.

Koelsch, S. (2011). Toward a neural basis of music perception - a review and updated model. *Frontiers in Psychology, 2*(110). https://doi.org/10.3389/fpsyg.2011.00110.

Kohler, E., Keysers, C., Umiltà, M. A., Fogassi, L., Gallese, V., & Rizzolatti, G. (2002). Hearing sounds, understanding actions: action representation in mirror neurons. *Science, 297*, 846–848.

Lahav, A., Boulanger, A., Schlaug, G., & Saltzman, E. (2005). The power of listening: auditory-motor interactions in musical training. *Annals of the New York Academy of Sciences, 1060*(1), 189–194.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: audiomotor recognition network while listening to newly acquired actions. *The Journal of Neuroscience, 27*(2), 308–314.

Lahav, A., Katz, T., Chess, R., & Saltzman, E. (2013). Improved motor sequence retention by motionless listening. *Psychological Research, 77*(3), 310–319.

Langers, D. R., & Melcher, J. R. (2011). Hearing without listening: functional connectivity reveals the engagement of multiple nonauditory networks during basic sound processing. *Brain Connectivity, 1*(3), 233–244.

Lanzilotto, M., Perciavalle, V., & Lucchetti, C. (2013). Auditory and visual systems organization in Brodmann Area 8 for gaze-shift control: where we do not see, we can hear. *Frontiers in Behavioral Neuroscience, 7*(198). https://doi.org/10.3389/fnbeh.2013.00198.

Lepage, J.-F., Tremblay, S., Nguyen, D. K., Champoux, F., Lassonde, M., & Théoret, H. (2010). Action related sounds induce early and late modulations of motor cortex activity. *NeuroReport, 21*, 250–253.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition, 21*(1), 1–36.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of speech code. *Psychological Review, 74*, 431–461.

Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of supplementary motor areas in auditory processing and auditory imagery. *Trends in Neurosciences, 39*(8), 527–542.

Luczak, A., Barthó, P., & Harris, K. D. (2009). Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron, 62*, 413–425.

Luo, C., Guo, Z. W., Lai, Y. X., Liao, W., Liu, Q., Kendrick, K. M., Yao, D. Z., & Li, H. (2012). Musical training induces functional plasticity in perceptual and motor networks: Insights from resting-state fMRI. *PLoS One, 7*(5), e36568. https://doi.org/10.1371/journal.pone.0036568.

Lyon, C. (2014). Beyond vision: Extending the scope of a sensorimotor account of perception. In J. M. Bishop & A. O. Martin (Eds.), *Contemporary sensorimotor theory* (pp. 127–136). Switzerland: Springer.

Manley, G. A., Fay, R. R., & Popper, A. N. (Eds.). (2008). *Active processes and Otoacoustic emissions in hearing*. New York: Springer.

Maoiléidigh, D. Ó., & Ricci, A. J. (2019). A bundle of mechanisms: Inner-ear hair-cell mechanotransduction. *Trends in Neurosciences, 42*(3), 221–236.

Martin, P. (2008). Active hair-bundle motility of the hair cells of vestibular and auditory organs. In G. A. Manley, R. R. Fay, & A. N. Popper (Eds.), *Active processes and Otoacoustic emissions in hearing* (pp. 93–143). New York: Springer.

Mathias, B., Palmer, C., Perrin, F., & Tillmann, B. (2015). Sensorimotor learning enhances expectations during auditory perception. *Cerebral Cortex, 25*(8), 2238–2254.

Mcdonald, J. S. P., & Lavie, N. (2011). Visual perceptual load induces inattentional deafness. *Attention, Perception, & Psychophysics, 73*(6), 1780–1789.

McGann, M., & De Jaegher, H. (2009). Self-other contingencies: enacting social perception. *Phenomenology and the Cognitive Sciences, 8*(4), 417–437.

Meister, I. G., Boroojerdi, B., Foltys, H., Sparing, R., Huber, W., & Töpper, R. (2003). Motor cortex hand area and speech: Implications for the development of language. *Neuropsychologia, 41*(4), 401–406.

Molloy, K., Griffiths, T. D., Chait, M., & Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. *The Journal of Neuroscience, 35*(49), 16046–16054.

Moray, N. (1959). Attention in dichotic listening: affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11*(1), 56–60.

Möttönen, R., & Watkins, K. E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *The Journal of Neuroscience, 29*(31), 9819–9825.

Myin, E. (2016). Perception as something we do. *Journal of Consciousness Studies, 23*(5–6), 80–104.

Naumann, M., Magyar-Lehmann, S., Reiners, K., Erbguth, F., & Leenders, K. L. (2000). Sensory tricks in cervical dystonia: Perceptual dysbalance of parietal cortex modulates frontal motor programming. *Annals of Neurology, 47*, 322–328.

Noë, A. (2002). Is the visual world a grand illusion? *Journal of Consciousness Studies, 9*(5–6), 1–12.

Noë, A. (2004). *Action in perception*. Cambridge: The MIT Press.

Noë, A. (2009). *Out of our heads: Why you are not your brain, and other lessons from the biology of consciousness*. New York: Hill and Wang.

O'Regan, J. K. (2011). *Why red Doesn't sound like a bell: Understanding the feel of consciousness*. New York: Oxford University Press.

O'Regan, J. K., Rensink, R. A., & Clark, J. J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature, 398*, 34.

O'Regan, J. K., Deubel, H., Clark, J. J., & Rensink, R. A. (2000). Picture changes during blinks: Looking without seeing and seeing without looking. *Visual Cognition, 7*(1/2/3), 191–211.

O'Regan, J. K., Myin, E., & Noë, A. (2005). Sensory consciousness explained (better) in terms of 'corporality' and 'alerting capacity'. *Phenomenology and the Cognitive Sciences, 4*, 369–387.

Oh, S.-Y., Boegle, R., Ertl, M., Stephan, T., & Dieterich, M. (2018). Multisensory vestibular, vestibular-auditory, and auditory network effects revealed by parametric sound pressure stimulation. *Neuroimage, 176*, 354–363.

O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences, 24*(5), 939–1031.

Palomar-García, M. Á., Zatorre, R. J., Ventura-Campos, N., Bueichekú, E., & Ávila, C. (2017). Modulation of functional connectivity in auditory-motor networks in musicians compared with nonmusicians. *Cerebral Cortex, 27*(5), 2768–2778.

Pascal, F., & O'Regan, K. (2008). Commentary on Mossio and Taraborelli: is the enactive approach really sensorimotor? *Consciousness and Cognition, 17*(4), 1341–1342.

Pazzaglia, M., Pizzamiglio, L., Pes, E., & Aglioti, S. M. (2008). The sound of actions in apraxia. *Current Biology, 18*, 1766–1772.

Perrett, S., & Noble, W. (1997). The contribution of head motion cues to localization of low-pass noise. *Perception & Psychophysics, 59*(7), 1018–1026.

Pessoa, L., Thompson, E., & Noë, A. (1998). Finding out about filling-in: a guide to perceptual completion for visual science and the philosophy of perception. *Behavioral and Brain Sciences, 21*(6), 723–802.

Phillips-Silver, J., & Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain and Cognition, 67*(1), 94–102.

Pizzamiglio, L., Aprile, T., Spitoni, G., Pitzalis, S., Bates, E., D'Amico, S., & Di Russo, F. (2005). Separate neural systems for processing action- or non-action-related sounds. *Neuroimage, 24*, 852–861.

Pulvermüller, F., Huss, M., Kherif, F., del Prado Martin, F. M., Hauk, O., & Shtyrov, Y. (2006). Motor cortex maps articulatory features of speech sounds. *Proceedings of the National Academy of Sciences of the USA, 103*(20), 7865–7870.

Rauschecker, J. P. (1998). Cortical processing of complex sounds. *Current Opinion in Neurobiology, 8*(4), 516–521.

Rauschecker, J. P. (2018). Where, when, and how: are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex, 98*, 262–268.

Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience, 12*(6), 718–724.

Raveh, D., & Lavie, N. (2015). Load-induced inattentional deafness. *Attention, Perception, & Psychophysics, 77*(2), 483–492.

Repp, B. H. (1992). Perceptual restoration of a "missing" speech sound: auditory induction or illusion? *Perception & Psychophysics, 51*(1), 14–32.

Riecke, L., van Opstal, A. J., Goebel, R., & Formisano, E. (2007). Hearing illusory sounds in noise: sensory-perceptual transformations in primary auditory cortex. *The Journal of Neuroscience, 27*(46), 12684–12689.

Riecke, L., van Opstal, A. J., & Formisano, E. (2008). The auditory continuity illusion: a parametric investigation and filter model. *Perception & Psychophysics, 70*(1), 1–12.

Rietveld, E., Denys, D., & Van Westen, M. (2018). Ecological-enactive cognition as engaging with a field of relevant affordances: The skilled intentionality framework (SIF). In A. Newell, L. De Bruin, & S. Gallagher (Eds.), *The Oxford handbook of 4E cognition* (pp. 41–70). Oxford: Oxford University Press.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences, 21*, 188–194.

Rizzolatti, G., & Sinigaglia, C. (2016). The mirror mechanism: a basic principle of brain function. *Nature Reviews Neuroscience, 17*, 757–765.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131–141.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews Neuroscience, 2*, 661–670.

Rosenblum, L. D., Gordon, M. S., & Jarquin, L. (2000). Echolocating distance by moving and stationary listeners. *Ecological Psychology, 12*(3), 181–206.

Ross, J. M., Iversen, J. R., & Balasubramaniam, R. (2016). Motor simulation theories of musical beat perception. *Neurocase, 22*(6), 558–565.

Schalles, M. D., & Pineda, J. A. (2015). Musical sequence learning and EEG correlates of audiomotor processing. *Behavioural Neurology, 2015*(638202), 1–11. https://doi.org/10.1155/2015/638202.

Schomers, M. R., & Pulvermüller, F. (2016). Is the sensorimotor cortex relevant for speech perception and understanding? An integrative review. *Frontiers in Human Neuroscience, 10*(435). https://doi.org/10.3389/fnhum.2016.00435.

Schumann, F., & O'Regan, J. K. (2017). Sensory augmentation: integration of an auditory compass signal into human perception of space. *Scientific Reports, 7*(42197). https://doi.org/10.1038/srep42197.

Seth, A. K. (2014). A predictive processing theory of sensorimotor contingencies: explaining the puzzle of perceptual presence and its absence in synaesthesia. *Cognitive Neuroscience, 5*(2), 97–118. https://doi.org/10.1080/17588928.2013.877880.

Shahin, A. J., Bishop, C. W., & Miller, L. M. (2009). Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage, 44*, 1133–1143.

Silva Pereira, C., Teixeira, J., Figueiredo, P., Xavier, J., Castro, S. L., & Brattico, E. (2011). Music and emotions in the brain: familiarity matters. *PLoS One, 6*(11), e27241. https://doi.org/10.1371/journal.pone.0027241.

Silverman, D. (2013). Sensorimotor enactivism and temporal experience. *Adaptive Behavior, 21*(3), 151–158.

Silverman, D. (2018). Bodily skill and internal representation in sensorimotor perception. *Phenomenology and the Cognitive Sciences, 17*, 157–173.

Simmons, F. B. (1964). Perceptual theories of middle ear muscle function. *Annals of Otology, Rhinology and Laryngology, 73*, 724–739.

Skipper, J. I., Devlin, J. T., & Lametti, D. R. (2017). The hearing ear is always found close to the speaking tongue: review of the role of the motor system in speech perception. *Brain and Language, 164*, 77–105.

Steinbrink, C., Ackermann, H., Lachmann, T., & Riecker, A. (2009). Contribution of the anterior insula to temporal auditory processing deficits in developmental dyslexia. *Human Brain Mapping, 30*, 2401–2411.

Stephan, M. A., Lega, C., & Penhune, V. B. (2018). Auditory prediction cues motor preparation in the absence of movements. *Neuroimage, 174*, 288–296.

Stupacher, J., Hove, M. J., Novembre, G., Schütz-Bosbach, S., & Keller, P. E. (2013). Musical groove modulates motor cortex excitability: a TMS investigation. *Brain and Cognition, 82*(2), 127–136.

Thompson, E. (2007). *Mind in life: Biology, phenomenology, and the sciences of mind*. Cambridge: Harvard University Press.

Ticini, L. F., Schütz-Bosbach, S., Weiss, C., Casile, A., & Waszak, F. (2012). When sounds become actions: higher-order representation of newly learned action sounds in the human motor system. *Journal of Cognitive Neuroscience, 24*(2), 464–474.

Ticini, L. F., Schütz-Bosbach, S., & Waszak, F. (2017). Mirror and (absence of) counter-mirror responses to action sounds measured with TMS. *Social Cognitive and Affective Neuroscience, 12*(11), 1748–1757.

Ticini, L. F., Schütz-Bosbach, S., & Waszak, F. (2019). From goals to muscles: motor familiarity shapes the representation of action-related sounds in the human motor system. *Cognitive Neuroscience, 10*(1), 20–29.

Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition, 97*(2), B25–B34.

Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: cognitive science and human experience*. Cambridge: MIT Press.

Vörös, S., Froese, T., & Riegler, A. (2016). Epistemological odyssey: introduction to special issue on the diversity of enactivism and neurophenomenology. *Constructivist Foundations, 11*(2), 189–203.

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology, 27*(4), 339–368.

Ward, D., Silverman, D., & Villalobos, M. (2017). Introduction: the varieties of enactivism. *Topoi, 36*(3), 365–375.

Warren, R. M., Obusek, C. J., & Ackroff, J. M. (1972). Auditory induction: perceptual synthesis of absent sounds. *Science, 176*, 1149–1151.

Watkins, K. E., Strafella, A. P., & Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia, 41*(8), 989–994.

Wayland, J. F., Levin, D. T., & Varakin, D. A. (2005). Inattentional blindness for a noxious multimodal stimulus. *American Journal of Psychology, 118*(3), 339–352.

Wilson, S. M., & Iacoboni, M. (2006). Neural responses to non-native phonemes varying in producibility: evidence for the sensorimotor nature of speech perception. *Neuroimage, 33*(1), 316–325.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*(9), 338–347.

Wood, N., & Cowan, N. (1995). The cocktail party phenomenon revisited: how frequent are attention shifts to one's name in an irrelevant auditory channel? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 255–260.

Woods, E. A., Hernandez, A. E., Wagner, V. E., & Beilock, S. L. (2014). Expert athletes activate somatosensory and motor planning regions of the brain when passively listening to familiar sports sounds. *Brain and Cognition, 87*, 122–133.

Affiliations

Tom Froese[1,2] · Ximena González-Grandón[1,3]

[1]   Institute for Applied Mathematics and Systems Research (IIMAS), National Autonomous University of Mexico (UNAM), Circuito Escolar (IIMAS-UNAM), Ciudad Universitaria, 04510 Mexico City, Mexico

[2]   Center for the Sciences of Complexity (C3), UNAM, Mexico City, Mexico

[3]   Institute of Philosophy and Sciences of Complexity (IFICC), Santiago, Chile