

Is song processing distinct and special in auditory cortex?

Ilana Harris¹, Efe C. Niven¹, Alex Griffin², Sophie K Scott^{1†}

¹Institute of Cognitive Neuroscience, University College London, London, UK

²Department of Psychology, University of Cambridge, Cambridge, UK

†email: sophie.scott@ucl.ac.uk

ORCID IDs:

- Ilana: 0000-0002-8085-5628
- Efe: 0009-0008-2687-9013
- Alex: 0009-0003-5509-4104
- Sophie: 0000-0001-7510-6297

Abstract | Is the singing voice processed distinctively in the human brain? In this Perspective, we discuss what might distinguish song processing from speech processing in the light of recent work suggesting that some cortical neuronal populations respond selectively to song and outline the implications for our understanding of auditory processing. We review the literature regarding the neural and physiological mechanisms of song production and perception and show that this provides evidence for key differences between song and speech processing. We conclude by discussing the significance of the notion that song processing is special in terms of how this might contribute to theories of the neurobiological origins of vocal communication and to our understanding of the neural circuitry underlying sound processing in the human cortex.

[H1] Introduction

Song is a distinct kind of human vocalization. From an acoustic perspective, the sequence of syllables of which song is composed are timed in a structured rhythmic fashion (where time intervals between consecutive events are generally equal) and are sung to a clear and formal pitch profile (tune, or melody), facilitating perceptual differentiation of song from speech. By contrast, speech (the other major type of human vocalization) has a weaker sense of pitch, a less structured intonation and (typically) an anisochronous rhythm (FIG. 1)^{1,2,3,4}. From a production perspective, human song and speech are produced via the same vocal tract system,

under voluntary control; however, there are some differences in the neural systems recruited for their production^{5,6}. From a perception perspective, there is considerable overlap between the brain areas activated by song and speech; however, several key differences do exist. Song —like music more generally — seems to preferentially evoke activity in brain networks involved in reward processing⁷, driving pleasurable experiences and highly motivating engagement with music⁸. Song perception also places higher demands on the brain areas involved in pitch processing than speech^{9,10}. In addition, there are differences in the sensorimotor processes involved in the acquisition of song and speech^{11,12,9}. A recent study using electrophysiological measures has shown evidence for song selective processing in auditory cortical fields (over and above activity elicited by instrumental music and speech)¹³. Although it is not yet clear when in brain development this song-selectivity arises, we do know that song is salient and important for communicating emotion between mother and child from a very young age: newborns display a preference for infant-directed song over adult-directed song^{14,15}, even when the infants are born of deaf mothers¹⁶.

The human capacity for song is universal^{17,18} and may be very evolutionarily ancient: it has been argued that Neanderthals communicated via a song-like protolanguage of musical emotive expression¹⁹ (but see REFs^{20,21} for the alternative argument that music predated language). Unlike speech, which is primarily used for interactions and communication, song is key for transmitting more semantically ambiguous socio-emotional information^{22–24}. For example, song is important for social bonding²⁵ and a link between song forms and unique socio-emotional functions has been suggested to exist in many cultures (for example, lullabies across cultures tend to have a simple structure and falling contours, which may enable them to modulate arousal)^{26–29}. There are also universal features of song perception across cultures, and these are likely to be influenced by biological and cognitive constraints: for example, rhythmic expectations are biased towards simple integer ratios for tapping and singing, and the tendency to expect certain rhythms over others has been theorized to be underscored by recurrent attractor networks³⁰ and oscillatory dynamics^{31–33} governing brain activity. These regular rhythms contrast to the rhythms typically used in speech, which are typically less

regular, and are influenced by structural aspects of spoken languages^{23,34}. Similarly, pitch perception is limited by a compressed representation of higher frequencies, and a deterioration of the capacity of the brain to extract pitch for extremely high or low frequencies³⁵.

The primary aim of this Perspective article is to review the physiological and neural mechanisms of song production and perception and to highlight how the human brain and body might be adapted for song. We also elucidate key differences between song and speech processing, including the recent work suggesting that some cortical neuronal populations respond selectively to song. We contextualize this discussion with suggestions for further research into evolutionary and developmental aspects of human vocal communication and the neural circuitry for sound processing in human cortex. We restrict the scope of our article to song and its direct comparisons with speech, but do not comprehensively examine the similarities between music and language (see REF³⁶ for an extensive review and REF³⁷ for a review of neural overlap between these two domains).

[H1] Mechanisms of song production

It is generally accepted that the auditory circuits underpinning vocal productive and perceptive mechanisms are linked, particularly for speech and song³⁸. In the following section, we examine the physiological and neural mechanisms of song production and investigate how the human brain and body may have selectively adapted to support song production.

[H2] Physiological bases of song production

All vocalizations can be analyzed from a production perspective by using an important framework that is denoted the source-filter theory. Vibrations of the vocal folds in the larynx give rise to a glottal wave, referred to as the 'source signal'^{39,40}. This source signal is then 'filtered' by the shape/length of the supralaryngeal vocal tract and movements of supralaryngeal articulatory structures, which change the spectrotemporal structure of the glottal wave and give rise to the plethora of vocal sounds that can be produced⁴¹⁻⁴³. The ability

to sustain controlled breathing and vibration of vocal chords in the larynx while dynamically changing articulatory filters allows for the production of complex human voluntary vocalizations^{39,41,44} (FIG. 1). In particular, the two-tube configuration of the human supralaryngeal vocal tract and the globular tongue have been proposed to support greater anatomical efficiency of sequential and dynamic modifications of the supralaryngeal vocal tract, which are needed for producing complex vocalizations inherent to speech and song⁴⁵ (FIG. 2A).

However, the vocal fold vibrations that underlie song tend to be more stable in frequency than those that enable speech⁴⁶, whereas speech's dynamic filtering tends to occur at faster rates than that of song⁴⁷. The vocal fold vibrations and filtering mechanics used to produce song (FIG. 2A) also helps to distinguish it from other melodic sounds produced by the vocal apparatus: whereas whistling and humming both contain melody, only humming involves vocal fold vibrations, and neither incorporates the array of rapid, sequential, and detailed movements of the articulators that give rise to the complex and continuous filters that distinguish song and speech^{48,49}.

[H2] Neural Bases of Song Production

Voluntary control of phonated sounds (such as speech and song) appears to be a uniquely human adaptation. Great apes can control non-phonated oral sounds (such as whistles) produced by the supralaryngeal vocal tract⁵⁰, but do not appear to be able to voluntarily modify phonated sounds (as would be necessary for vocal learning). Thus, researchers have sought to understand the neurobiological basis of this human-specific capacity. Within the human cortex, there are direct projections from the lateral motor cortex to laryngeal motor neurons in the nucleus ambiguus which underpin fine motor control of the larynx (FIG. 2B)^{51,52}. These direct projections have long been hypothesized to underpin the capacity to produce uniquely human vocalizations; however, homologous direct projections in the macaque and mice brains (from primary motor cortex to vocal motor neurons) that exert cortical control over vocal pitch have recently been found⁵³. Similarly, diffusion tractography studies have shown that the laryngeal

motor cortex networks (located within the lateral motor cortex) in non-human primates and humans are similar, although there are many more projections from the laryngeal motor cortex to somatosensory and inferior parietal cortex in humans^{54,55}. Very recent research conducted on over 12,000 individuals residing in Iceland suggests that the vocal pitch and vowel acoustics that an individual can produce are influenced by their genes, with vocal pitch being correlated with common variants in *ABCC9*, a gene encoding a potassium channel subunit⁵⁶. Follow-up comparative studies in other mammals would be extremely useful to determine how genetic determinants of vocal pitch and/or quantitative differences in effective connectivity support the fine-grained volitional control of vocal pitch underpinning the capacity for human song.

Learned (voluntary) vocalizations are thought to be elicited via the volitional vocalization network (VVN), which includes the bilateral primary motor and somatosensory cortex, the supplementary motor area, the auditory cortex, subcortical structures, and the prefrontal and insular cortices⁵⁷ (FIG. 2B; though it should be noted that an alternative ‘single vocal system’ model proposes that there is a common sensorimotor vocal system for speech, emotional vocalizations, and song^{11,58}). The cortical auditory processing that is important for song production mainly occurs in the superior temporal gyrus (STG) and/or the superior temporal sulcus (STS), where information — including pitch (encoded in the right lateral Heschl’s gyrus), pitch changes (encoded in the right STG, planum polare and planum temporale), melodic phrases (encoded in the bilateral STG), and vocal timbre (encoded in the STS, with functional divides among its anterior, mid/anterior, and posterior portions) — informs subsequent motor adjustments via the insula^{59–61}. Somatosensory processing in somatosensory cortex is also particularly important for vocal pitch regulation and seems to dominate the sensory inputs guiding trained singers’ vocal tract adjustments. Indeed, when compared with untrained singers, classically trained singers show structural differences in somatosensory regions, including increased right-hemisphere gray matter volume in the primary (S1) and secondary (S2) somatosensory cortex⁶² and greater functional connectivity between the anterior insula and the regions of sensorimotor cortex representing the larynx and diaphragm⁶³. Furthermore, a

functional imaging study found increased BOLD responses in regions of S1 receiving laryngeal and mouth input during singing when comparing opera singers to untrained singers⁶⁴.

The left anterior insula has been shown to underlie the coordination of speech articulation^{65,66}, while the right anterior insula is important for integrating somatosensory and auditory perceptual information during singing^{67,68} and has been shown to support vocal tract movements during singing^{69,70} and during melodic elements of non-song vocalizations⁷¹. One study found that singers showed decreased activation of the right anterior insula when singing with anesthetized vocal folds, suggesting that information about laryngeal movement was not integrated into subsequent motor commands when it was deemed unreliable⁶⁷. However, in a second study in which auditory perception was masked with loud noise, trained singers maintained accurate pitch and showed enhanced activation in their anterior insulas and intraparietal lobules. This contrasted with untrained singers (who also had significantly decreased pitch accuracy), indicating trained singers' greater use of somatosensory information to maintain accurate pitch, as noted above⁶⁸. Furthermore, repetitive transcranial magnetic stimulation (TMS) applied to the part of S1 receiving input from the larynx has been shown to enhance pitch accuracy and pitch stability in untrained singers, suggesting an experience-dependent role of somatosensory integration during singing⁷².

Recent accounts have revealed human volitional vocal control mechanisms that relate to the production of information in speech beyond the linguistic message, such as the vocal expression of identity (such as that often expressed via song)⁵. For example, normal speech production recruits the left posterior frontal gyrus⁵⁷, while bilateral frontal gyri are recruited to a greater extent when talkers produce speech with different accents or imitate a particular individual, relative to normal speech⁷³. Bilateral inferior frontal gyri (IFG) versus right-lateralized IFG activation has also been reported for song versus speech^{48,69}. Right-lateralized IFG responses are seen when talkers align the rhythm and melody of their voice with that of another when speaking in unison⁷⁴. Notably, one study reported suppression of activity in the

right IFG during normal speech production⁷⁵, suggesting that there is considerable complexity in how this frontal network is recruited in speech and song production.

[H1] Mechanisms of song perception

Some general auditory perceptual processing constraints apply to song, as to any sound: sounds only have structure as they evolve over time, and only exist as result of physical actions in the world. Sounds need to be processed fast, and their spectro-temporal structure includes information about their meaning as well as the means of production⁷⁶. Auditory processing in the human cortex has been theorized to follow a dual-stream framework, similar to that proposed for the visual domain, in which a rostral 'what' pathway supports auditory object identification⁷⁷ and caudal 'where/how' pathways support the representation of auditory space and action-guiding computations⁷⁷.

Speech and song evoke similar perceptual processes and corresponding cortical activity along these what and where/how pathways (FIG. 3A). A study that identified common patterns of activation for the perception of speech and music reveals bilateral patterns of activation in the dorsolateral temporal lobes⁷⁸; furthermore, unlike instrumental music, the perception of both speaking and the singing voice result in extensive activation of the bilateral STS⁷⁸. However, there are some distinct differences in within this bilateral pattern: within the rostral pathway, activity is relatively left-lateralized for phoneme discrimination⁷⁹ and intelligibility⁸⁰ and relatively right-lateralized for chord discrimination⁸¹ and normal speech intonation⁸². The mechanisms underlying these hemispheric asymmetries are yet to be determined^{83,84}.

The caudal pathways (FIG. 3A) underpin sensorimotor integration^{77,85-87}, which is important for the vocal learning necessary to perform speech and song, and for the guidance of the production of speech and song^{88,89}. Caudal auditory fields have fine temporal acuity, underscoring their role in performing action-guiding computations⁹⁰. A region within the caudal medial auditory cortex shows activity during the perception and covert (imagined) production

of both song and speech⁹¹. This field has been hypothesized to support domain-general auditory-motor integration, but also to contain subsidiary regions that are more responsive towards domain-specific sound categories (that is, regions more selective to speech than music, and vice versa). Understanding the influence of sensorimotor loops on song perception has been particularly important for better understanding previously conceived universal properties of the music perceptual system, in which what have been called the “universal physical constraints of [musical] performance” give rise to certain melodic regularities observed in music and consequently dictate constraints on the perception of melody^{92,93}. For example, the tendency of melodies to descend following a large ascending interval (referred to as a ‘post-skip reversal’) can be explained by the fact that accurately singing a large ascending interval is relatively difficult and exercises limitations on singers’ vocal ranges, so singers will be more likely to consequently descend rather than continue to ascend after such a large ascending interval to ensure that the sung melody can stay within their vocal range^{94,95}. Studies have shown that the constraints imposed by a singer’s vocal range can be used to explain the observed regularity of post-skip reversals in music⁹³.

[H2] Right-hemispheric lateralization for song?

As in the brain networks involved in song production, there is some degree of hemispheric lateralization in the auditory fields responsible for song perception, in terms of their responses to auditory stimuli. Certain domain-general properties of acoustic signals that are particularly relevant for song processing tend to be processed in the right hemisphere of the brain⁸³. In primary auditory cortex, pitch extraction occurs bilaterally in Heschl’s gyrus⁶¹, but is right-lateralized for ‘missing fundamental’ sounds, that is, a sound whose harmonic structure implies a particular fundamental frequency, but no sound at that frequency is actually present⁹⁶ (see **Box 1**). Furthermore, a relative right hemispheric preference has been observed for the perception of pitch variation^{97–99}. The computation of pitch distance between consecutive tones (that is, of intervals) activates the right planum temporale⁵⁹ and higher-level processing of pitch patterns (across a melodic sequence, for example) activates a more broadly distributed

right-lateralized fronto-parietal network¹⁰⁰. Song frequently involves the production of longer syllables (Fig 1) and the perceptual processing of relatively longer sounds (hundreds of milliseconds, rather than tens of milliseconds) is also right-lateralized⁸³. Overlapping frontal regions in the right inferior frontal gyrus (IFG) support processing of emotional prosody and musical syntax¹⁰¹.

In line with these observations on the processing of acoustic information more generally, a right-lateralized network involved in song processing has been identified that includes primary auditory regions and extends to the higher-level auditory association areas (i.e., in planum temporale) and frontal cortex (i.e., IFG)^{102–105}. This observed lateralization for song processing seems to be particularly important for melodic perception: in an fMRI study investigating melody recognition and sentence recognition for 100 acapella songs, the neural decoding of melody was localized to the right lateral anterior STG, whereas the neural decoding of sentences was localized to the left lateral anterior STG¹⁰⁶. It was also found that perception of melodies in spoken poem stimuli elicited activity in temporal but not frontal regions involved in the right-lateralized melody network, suggesting that information processing of higher-order melody (that is, melody that is dependent on musical syntax) is more evident for song than for poems¹⁰.

It has been suggested that song processing places greater demands on sensorimotor-related auditory processing than speech⁹. Explicit processing of isochronous rhythm in speech — similar to the rhythmic properties of song — has been shown to specifically activate the right auditory cortex (posterior STG¹⁰⁷), although bilateral activity is generally found for musical rhythms (within a cortico-subcortical network¹⁰⁸) and for beat-inducing metric simple rhythms (within the putamen and STG¹⁰⁹). Nevertheless, music and language both activate bilateral auditory processing networks when compared to baseline, with reported lateralization often resulting from direct comparison of activation patterns for speech vs. song (see, for example, REF¹⁰).

[H2] Voice and music selectivity in auditory cortex

Selectivity for the perception of specific conspecific vocalizations has been shown in single-unit recordings from neurons along the lateral auditory belt in the macaque monkey^{110,111}. Similarly, in humans, temporal voice areas (TVAs), defined by identifying brain regions that respond to a variety of different human vocalizations have been described bilaterally along the upper bank of the STS: one is located anteriorly, near the temporal pole; one centrally, near anterior Heschl's gyrus; and one posterior to Heschl's gyrus, extending rostrally and caudally to the planum temporale^{112,113}. Later work showed that the right TVAs showed stronger activation for sung vowels versus harmonic-to-noise matched instrumental sounds and for vocal stimuli versus a baseline¹¹⁴. Bilateral clusters in the prefrontal cortex (IFG) which display voice sensitivity and connectivity with TVAs in auditory cortex have also been identified, and may contribute to identity- and emotion-related vocal processing¹¹⁵⁻¹¹⁷. TVAs emerge early in life and homologous regions have been identified in macaques^{110,111,118}, suggesting that the voice's salience is evolutionarily old and developmentally important. Indeed, given that both faces and voices can provide key information about the identity of the speaker, voices have been proposed as being the equivalent to faces in the auditory domain (see **Box 2**).

In human auditory cortex, music-selectivity has been found in regions posterior and anterior to the primary auditory cortex¹¹⁹⁻¹²¹, whereas speech-selectivity is found in more medial non-primary auditory regions^{80,121,122}. Importantly, these category-selective regions show insensitivity to spectrotemporally matched synthetic sounds and non-linear cortical responses, in contrast to spectrotemporally selective regions in primary auditory cortex that show linear responsivity to spectrographic features^{121,123}. A comparative fMRI study found clusters of voxels selective to particular harmonic tones (but not to noisy sounds with no harmonic structure) in humans in the primary auditory cortex and extending anteriorly into non-primary regions, whereas macaques showed significant noise-selective, but not tone-selective, clusters¹²⁴. This suggests that harmonicity is an important low-level feature for human audition and agrees with pitch being a perceptual property of both speech and music.

A recent study reported the discovery of a population of neurons in the bilateral auditory cortex that are selective for song^{13,125} (FIG. 3B), relative to speech, music, or speech plus music. This study used electrocorticography (ECoG), in which electric potentials are directly recorded from an exposed area of the cortex in individuals receiving presurgical treatments¹²⁶. This method allows for heightened spatial and temporal specificity in comparison to fMRI and is particularly useful for isolating discrete neuronal responses in the auditory cortex. The sounds presented were short sequences of a variety of different kinds of sound, including speech, instrumental music, and songs. The responses of neuronal populations were analyzed via component analysis and were assessed using both a data-driven approach, which did not pre-define sound categories, and a hypothesis-driven approach, which involved pre-defined sound categories. A non-linear response to song that was supra-additive to activity elicited by instrumental music and speech and that far exceeded activity for spectro-temporally matched synthetic sounds was discovered in a population localized to middle and anterior STG. Voxel-based component modeling was used to further support these ECoG results, showing that song selectivity was evident, albeit weaker, in previously gathered fMRI data. This strongly suggests that there are auditory fields in humans — found bilaterally and near to music- and speech-selective areas in the middle and anterior STG — that show selective responses to song.

[H2] The implications of song selective responses

The discovery of song selective cortical responses is important, not least because the additive effects of speech and music suggests that the responses are not merely to ‘musical speech’. The early preference of infants for song may well be relevant to the development of such a song specific response, which in turn may be based on auditory experience of the low frequency pitch and rhythm of the sound of the mother’s voice in utero¹²⁷. Hearing develops in the last trimester of pregnancy and babies are born with some knowledge of sounds, with prenatal experience resulting in preference for the maternal voice within 2 days after birth¹²⁷, neonates’ increased neuronal response to native vs. foreign or flattened accents (in left-lateralized frontal, middle and superior temporal, central, and parietal areas¹²⁸), and the positive impact of

prosodic pitch variation on memory of words' phonetic properties within the first 2 months of life (known as 'prosodic bootstrapping'¹²⁹). During development, babies' preference for the singing voice may also relate to the way that they use the melody of speech to help segment out words from the constant stream of speech that they hear¹²⁹. The singing voice may thus form a scaffold for the acquisition of spoken language, as well as an immersion into musical structure and form. This early engagement with song may also have an important role in our emotional reactions to song and to music⁷.

A recent preprint reported greater responses to instrumental music (that is, simple monophonic music) in one-month old newborns than to infant-directed maternal speech in the non-primary auditory cortex, controlling for spectrotemporal modulation¹³⁰. This suggests that music selectivity arises early in life and that music may be more salient than speech at early developmental stages, paralleling the influential finding that babies from congenitally deaf mothers display a preference for infant-directed singing (despite having not heard singing prenatally)¹⁶. However, the acoustic exposure in utero makes it hard to establish whether these findings are due to innate prepotent responses or due to experience. As has been similarly suggested regarding music-selectivity, the development of song-selectivity is likely to rely on a combination of neural specializations driven by progressive modularization of the human brain during development^{131,132}. These may be innate (that is, present before hearing any stimuli, including prenatally), experience-dependent (given that long-term experience critically impacts music and language processing; see REF¹³³ for a review) and experience-expectant (for example, specialized circuits are molded during a critical or sensitive period). For example, prenatal exposure to the maternal voice could predispose infants to song-relevant components of sound (such as lower frequencies), with subsequent learning of hierarchical structure resulting in perception of song as a discrete category¹³⁴. Several existing accounts have likened regions displaying song or music selectivity to visual word form areas^{13,132}, for which it has been shown that regions displaying selectivity for written words have colonized regions displaying selectivity for faces due to neuronal recycling (that is, visual neural systems that have adapted fine-grained specializations for face recognition have been re-used for similar processes in

literacy)¹³⁵. Perhaps the prepotent predisposition in many auditory fields is a preference for the singing voice, and subsequent acquisition of spoken language colonizes much of this system in early development. Incorporating a production-based developmental trajectory into the investigation of this idea would be important (that is, how increased motor control over articulatory structures across development differentially affects song versus speech processing; see REF^{136,137} for accounts regarding speech motor development). This links debates about the evolutionary origins of music, speech, and song to the neural recycling hypotheses, and which sound category (if any) has parasitized the other becomes a central question.

Understanding how differences in neuronal encoding properties may give rise to song selectivity is also likely to shed light on organization of auditory cortex more generally. Primary auditory neurons have been shown to encode spectrotemporal peculiarities of sung vowels, but to display comparable activation for synthesized artificial controls with similar low-level acoustic fluctuations¹¹⁴. Activity in song-selective neural populations occurs at longer latencies (≥ 200 ms) than these spectrotemporally sensitive neurons in primary auditory cortex (≤ 200 ms¹³). Longer onset responses are characteristic of neurons engaged in higher-order perceptual processing, which integrate incoming perceptual information with prior experience⁹⁰. A recent study using electrocorticography found that longer integration windows support category-selective encoding in neurons in non-primary regions, whereas shorter integration windows support cochleogram-based encoding in neurons in primary auditory regions¹³⁸. Integration windows in non-primary regions seem to be context-invariant (that is, they are insensitive to previous auditory segments), with no apparent difference between responses to randomly segmented vs. intact naturalistic sounds. Direct comparisons of song versus speech stimuli in similar paradigms could shed light on the degree to which spatial clustering and temporal dynamics is indicative of domain-general hierarchical encoding in auditory cortex.

Understanding song selectivity will also shed light on the spatial clustering of networks underpinning category selectivity in human cortex. The song-selective population was reported

to be selective to ‘music with singing’ (as opposed to a single singing voice) and may therefore colonize subregions of music-selective or voice-only-song-selective regions. It is possible that category-selectivity in STG shows a similar organizational principle to that of sensorimotor production-based encoding in the caudal medial auditory cortex (in which regions of broader selectivity contain subsidiary regions with increased domain specificity⁹¹). Future studies should investigate neural selectivity for stylistically diverse examples of singing that includes only a single voice (such as popular music including rap, or chanting), and compare the spatial responses of such a population to that which is selective for music with singing. This would be particularly interesting, given accounts that music-making is “a solo endeavor” in some traditional societies^{132,139}, and would shed light on the neural template for song (in line with neural templates for voice or face)¹⁴⁰.

Future studies should also investigate the contribution of figure-ground encoding (that is, the separation of an auditory object from its surroundings) in primary auditory cortex to higher-level neural selectivity for song; in REF¹³, song was always presented alongside musical accompaniment. In Western recorded music (as used in this study), the singing voice typically represents the dominant object in such complex musical scenes, the default figure against the ground. Further studies should look at other kinds of configurations, for example by parametrically varying the number of voices in a multi-voice piece. In turn, this will help to elucidate the role of multivariate coding in auditory cortex. A recent study in non-human primates found that the presence of non-harmonically defined auditory figures (that is, complex sounds lacking a pitch percept, such as the sound of a slamming door) was signaled by a rate code in primary auditory cortex¹⁴¹. Notably, multi-unit activity corresponding to figure-ground responses was not spatially clustered, and it was unclear whether the activity arose from a single neuron, a local circuit, or a distributed population. It would be interesting to see whether similar figure-ground responsive neurons exist in human primary auditory cortex and, if so, whether they are equally as responsive to non-harmonic and harmonic figures.

[H1] Conclusion

In this Perspective, we have shown that the singing voice is distinct not only in terms of its distinguishing acoustic and perceptual features but also in terms of how it is produced. Production–perception links, which underpin the dual stream models of speech processing, also exist for song and music. The sound of a singing voice is inextricably linked to the vocal mechanisms that produce that song. It is possible that song comprises vocal actions and sounds that are particularly salient, acquired early in development, and are highly useful for the developing human brain. Studying song will help us to understand how vocalizations relevant for our species' survival are encoded in human cortex, and future work — particularly developmental — will shed light on how the human brain becomes predisposed for certain vocal actions such as song.

References

1. Lindblom, B. & Sundberg, J. The human voice in speech and singing. in *Handbook of Acoustics* 669–712 (Springer-Verlag New York, 2007).
2. Gerhard, D. Pitch-based acoustic feature analysis for the discrimination of speech and monophonic singing. *Can. Acoust.* **30**, 152–153 (2002).
3. Albouy, P., Mehr, S. A., Hoyer, R. S., Ginzburg, J. & Zatorre, R. J. Spectro-temporal acoustical markers differentiate speech from song across cultures. 2023.01.29.526133 Preprint at <https://doi.org/10.1101/2023.01.29.526133> (2023).
4. Yu, C. Y., Cabildo, A., Grahn, J. A. & Vanden Bosch der Nederlanden, C. M. Perceived rhythmic regularity is greater for song than speech: examining acoustic correlates of rhythmic regularity in speech and song. *Front. Psychol.* **14**, 1167003 (2023).
5. Scott, S. K. The neural control of volitional vocal production—from speech to identity, from social meaning to song. *Philos. Trans. R. Soc. B Biol. Sci.* **377**, 20200395 (2022).
6. Zuk, J. M., Loui, P. & Guenther, F. H. Neural Control of Speaking and Singing: The DIVA Model for Singing. 30.
7. Blood, A. J. & Zatorre, R. J. Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *Proc. Natl. Acad. Sci.* **98**, 11818–11823 (2001).
8. Zatorre, R. J. Musical pleasure and reward: mechanisms and dysfunction. *Ann. N. Y. Acad. Sci.* **1337**, 202–211 (2015).

9. Tierney, A., Dick, F., Deutsch, D. & Sereno, M. Speech versus Song: Multiple Pitch-Sensitive Areas Revealed by a Naturally Occurring Musical Illusion. *Cereb. Cortex N. Y. NY* **23**, 249–254 (2013).
10. Scharinger, M., Knoop, C. A., Wagner, V. & Menninghaus, W. Neural processing of poems and songs is based on melodic properties. *NeuroImage* **257**, 119310 (2022).
11. Cohen, A., Levitin, D. & Kleber, B. Brain Mechanisms Underlying Singing. in *The Routledge Companion to Interdisciplinary Studies in Singing* 79–96 (2020).
doi:10.4324/9781315163734-6.
12. Peretz, I., Gagnon, L., Hebert, S. & Macoir, J. Singing in the Brain: Insights from Cognitive Neuropsychology. *Music Percept.* **21**, 373–390 (2004).
13. Norman-Haignere, S. V. *et al.* A neural population selective for song in human auditory cortex. *Curr. Biol.* S0960982222001312 (2022) doi:10.1016/j.cub.2022.01.069.
14. Trainor, L. J., Clark, E. D., Huntley, A. & Adams, B. A. The acoustic basis of preferences for infant-directed singing. *Infant Behav. Dev.* **20**, 383–396 (1997).
15. Trainor, L. J., Austin, C. M. & Desjardins, R. N. Is Infant-Directed Speech Prosody a Result of the Vocal Expression of Emotion? *Psychol. Sci.* **11**, 188–195 (2000).
16. Masataka, N. Preference for infant-directed singing in 2-day-old hearing infants of deaf parents. *Dev. Psychol.* **35**, 1001–1005 (1999).
17. Brown, S. & Jordania, J. Universals in the world’s musics. *Psychol. Music* **41**, 229–248 (2013).
18. Mehr, S. A. *et al.* Universality and diversity in human song. *Science* **366**, eaax0868 (2019).

19. Mithen, S. J. *The Singing Neanderthals: The Origins of Music, Language, Mind, and Body*. (Harvard University Press, 2006).
20. Brandt, A., Slevc, R. & Gebrian, M. Music and Early Language Acquisition. *Front. Psychol.* **3**, 327 (2012).
21. Musical protolanguage. in *The Evolution of Language* (ed. Fitch, W. T.) 466–507 (Cambridge University Press, 2010). doi:10.1017/CBO9780511817779.015.
22. Cross, I. Is music the most important thing we ever did? Music, development and evolution. in *Music, Mind and Science* (ed. Suk Won Yi) 10–39 (Seoul National University Press, 1999).
23. Haiduk, F. & Fitch, W. T. Understanding Design Features of Music and Language: The Choric/Dialogic Distinction. *Front. Psychol.* **13**, 786899 (2022).
24. Cross, I. Music and meaning, ambiguity and evolution. in *Musical Communication* (eds. Miell, D., MacDonald, R. & Hargreaves, D.) 27–43 (Oxford university press, 2005). doi:10.1093/acprof:oso/9780198529361.003.0002.
25. Savage, P. E. *et al.* Music as a coevolved system for social bonding. *Behav. Brain Sci.* **44**, e59 (2021).
26. Mehr, S. A., Singh, M., York, H., Glowacki, L. & Krasnow, M. M. Form and Function in Human Song. *Curr. Biol.* **28**, 356-368.e5 (2018).
27. Unyk, A. M., Trehub, S. E., Trainor, L. J. & Schellenberg, E. G. Lullabies and Simplicity: A Cross-Cultural Perspective. *Psychol. Music* **20**, 15–28 (1992).
28. Trehub, S. E., Unyk, A. M. & Trainor, L. J. Maternal singing in cross-cultural perspective. *Infant Behav. Dev.* **16**, 285–295 (1993).

29. Trehub, S. & Trainor, L. Singing to infants: Lullabies and play songs. *Adv. Infancy Res.* **12**, 43–77 (1998).
30. Desain, P. & Honing, H. The Quantization of Musical Time: A Connectionist Approach. *Comput. Music J.* **13**, 56–66 (1989).
31. Large, E. W. & Snyder, J. S. Pulse and meter as neural resonance. *Ann. N. Y. Acad. Sci.* **1169**, 46–57 (2009).
32. Peper, C. E., Beek, P. J. & van Wieringen, P. C. W. Multifrequency coordination in bimanual tapping: Asymmetrical coupling and signs of supercriticality. *J. Exp. Psychol. Hum. Percept. Perform.* **21**, 1117–1138 (1995).
33. Bååth, R., Lagerstedt, E. & Gärdenfors, P. An Oscillator Model of Categorical Rhythm Perception. *Proc. Annu. Meet. Cogn. Sci. Soc.* **35**, 1803–1808 (2013).
34. Jacoby, N. & McDermott, J. H. Integer Ratio Priors on Musical Rhythm Revealed Cross-culturally by Iterated Reproduction. *Curr. Biol.* **27**, 359–370 (2017).
35. Jacoby, N. *et al.* Universal and Non-universal Features of Musical Pitch Perception Revealed by Singing. *Curr. Biol.* **29**, 3229–3243.e12 (2019).
36. Patel, A. D. *Music, Language, and the Brain*. (Oxford University Press, 2010).
37. Peretz, I., Vuhan, D., Lagrois, M.-É. & Armony, J. L. Neural overlap in processing music and speech. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140090 (2015).
38. Zatorre, R. J., Chen, J. L. & Penhune, V. B. When the brain plays music: auditory–motor interactions in music perception and production. *Nat. Rev. Neurosci.* **8**, 547–558 (2007).

39. Fant, G. Acoustic Theory of Speech Production. in *Description and Analysis of Contemporary Standard Russian* (De Gruyter Mouton, 1971).
doi:10.1515/9783110873429.
40. Briefer, E. & McElligott, A. G. Indicators of age, body size and sex in goat kid calls revealed using the source–filter theory. *Appl. Anim. Behav. Sci.* **133**, 175–185 (2011).
41. Titze, I. R. *Principles of voice production*. (Prentice Hall, 1994).
42. Titze, I. R. Nonlinear source–filter coupling in phonation: Theory). *J. Acoust. Soc. Am.* **123**, 2733–2749 (2008).
43. Tokuda, I. The Source–Filter Theory of Speech. in *Oxford Research Encyclopedia of Linguistics* (2021). doi:10.1093/acrefore/9780199384655.013.894.
44. Taylor, A. M. & Reby, D. The contribution of source–filter theory to mammal vocal communication research. *J. Zool.* **280**, 221–236 (2010).
45. Lieberman, P. *The Biology and Evolution of Language*. (1984).
46. Zatorre, R. J. & Baum, S. R. Musical Melody and Speech Intonation: Singing a Different Tune. *PLOS Biol.* **10**, e1001372 (2012).
47. Ozaki, Y. *et al.* Globally, songs and instrumental melodies are slower, higher, and use more stable pitches than speech [Stage 2 Registered Report]. Preprint at <https://doi.org/10.31234/osf.io/jr9x7> (2022).
48. Ozdemir, E., Norton, A. & Schlaug, G. Shared and distinct neural correlates of singing and speaking. *NeuroImage* **33**, 628–635 (2006).
49. Belyk, M. *et al.* Human larynx motor cortices coordinate respiration for vocal-motor control. *NeuroImage* **239**, 118326 (2021).

50. Wich, S. A. *et al.* A case of spontaneous acquisition of a human sound by an orangutan. *Primates* **50**, 56–64 (2009).
51. Iwatsubo, T., Kuzuhara, S., Kanemitsu, A., Shimada, H. & Toyokura, Y. Corticofugal projections to the motor nuclei of the brainstem and spinal cord in humans. *Neurology* **40**, 309–309 (1990).
52. Kuypers, H. G. Corticobular connexions to the pons and lower brain-stem in man: an anatomical study. *Brain J. Neurol.* **81**, 364–388 (1958).
53. Arriaga, G., Zhou, E. P. & Jarvis, E. D. Of Mice, Birds, and Men: The Mouse Ultrasonic Song System Has Some Features Similar to Humans and Song-Learning Birds. *PLOS ONE* **7**, e46610 (2012).
54. Kumar, V., Croxson, P. L. & Simonyan, K. Structural Organization of the Laryngeal Motor Cortical Network and Its Implication for Evolution of Speech Production. *J. Neurosci. Off. J. Soc. Neurosci.* **36**, 4170–4181 (2016).
55. Rauschecker, J. P. Where did language come from? Precursor mechanisms in nonhuman primates. *Curr. Opin. Behav. Sci.* **21**, 195–204 (2018).
56. Gisladdottir, R. S. *et al.* Sequence variants affecting voice pitch in humans. *Sci. Adv.* **9**, eabq2969 (2023).
57. Blank, S. C., Scott, S. K., Murphy, K., Warburton, E. & Wise, R. J. S. Speech production: Wernicke, Broca and beyond. *Brain* **125**, 1829–1838 (2002).
58. Belyk, M. & Brown, S. The origins of the vocal brain in humans. *Neurosci. Biobehav. Rev.* **77**, 177–193 (2017).

59. Hyde, K. L., Peretz, I. & Zatorre, R. J. Evidence for the role of the right auditory cortex in fine pitch resolution. *Neuropsychologia* **46**, 632–639 (2008).
60. Zarate, J. M. The neural control of singing. *Front. Hum. Neurosci.* **7**, 237 (2013).
61. Patterson, R. D., Uppenkamp, S., Johnsrude, I. S. & Griffiths, T. D. The Processing of Temporal Pitch and Melody Information in Auditory Cortex. *Neuron* **36**, 767–776 (2002).
62. Kleber, B. *et al.* Voxel-based morphometry in opera singers: Increased gray-matter volume in right somatosensory and auditory cortices. *NeuroImage* **133**, 477–483 (2016).
63. Zamorano, A. M. *et al.* Singing training predicts increased insula connectivity with speech and respiratory sensorimotor areas at rest. *Brain Res.* **1813**, 148418 (2023).
64. Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J. & Lotze, M. The brain of opera singers: experience-dependent changes in functional activation. *Cereb. Cortex N. Y. N 1991* **20**, 1144–1152 (2010).
65. Dronkers, N. F. A new brain region for coordinating speech articulation. *Nature* **384**, 159–161 (1996).
66. Wise, R. J., Greene, J., Büchel, C. & Scott, S. K. Brain regions involved in articulation. *Lancet Lond. Engl.* **353**, 1057–1061 (1999).
67. Kleber, B., Zeitouni, A. G., Friberg, A. & Zatorre, R. J. Experience-dependent modulation of feedback integration during singing: Role of the right anterior insula. *J. Neurosci.* **33**, 6070–6080 (2013).
68. Kleber, B., Friberg, A., Zeitouni, A. & Zatorre, R. Experience-dependent modulation of right anterior insula and sensorimotor regions as a function of noise-masked auditory feedback in singers and nonsingers. *NeuroImage* **147**, 97–110 (2017).

69. Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G. & Grodd, W. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport* **11**, 1997–2000 (2000).
70. Ackermann, H. & Riecker, A. The contribution of the insula to motor aspects of speech production: a review and a hypothesis. *Brain Lang.* **89**, 320–328 (2004).
71. Oh, A., Duerden, E. G. & Pang, E. W. The role of the insula in speech and language processing. *Brain Lang.* **135**, 96–103 (2014).
72. Finkel, S. *et al.* Intermittent theta burst stimulation over right somatosensory larynx cortex enhances vocal pitch-regulation in nonsingers. *Hum. Brain Mapp.* **40**, 2174–2187 (2019).
73. McGettigan, C. *et al.* T’ain’t what you say, it’s the way that you say it--left insula and inferior frontal cortex work in interaction with superior temporal regions to control the performance of vocal impersonations. *J. Cogn. Neurosci.* **25**, 1875–1886 (2013).
74. Jasmin, K. M. *et al.* Cohesion and Joint Speech: Right Hemisphere Contributions to Synchronized Vocal Production. *J. Neurosci.* **36**, 4669–4680 (2016).
75. Blank, S. C., Bird, H., Turkheimer, F. & Wise, R. J. S. Speech production after stroke: the role of the right pars opercularis. *Ann. Neurol.* **54**, 310–320 (2003).
76. Scott, S. K. Auditory processing — speech, space and auditory objects. *Curr. Opin. Neurobiol.* **15**, 197–201 (2005).
77. Rauschecker, J. P. & Scott, S. K. Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nat. Neurosci.* **12**, 718–724 (2009).
78. Whitehead, J. C. & Armony, J. L. Singing in the brain: Neural representation of music and voice as revealed by fMRI. *Hum. Brain Mapp.* **39**, 4913–4924 (2018).

79. Agnew, Z. K., McGettigan, C. & Scott, S. K. Discriminating between auditory and motor cortical responses to speech and nonspeech mouth sounds. *J. Cogn. Neurosci.* **23**, 4038–4047 (2011).
80. Scott, S. K., Blank, C. C., Rosen, S. & Wise, R. J. Identification of a pathway for intelligible speech in the left temporal lobe. *Brain J. Neurol.* **123 Pt 12**, 2400–2406 (2000).
81. Klein, M. E. & Zatorre, R. J. A role for the right superior temporal sulcus in categorical perception of musical chords. *Neuropsychologia* **49**, 878–887 (2011).
82. Kyong, J. S. *et al.* Exploring the roles of spectral detail and intonation contour in speech intelligibility: an fMRI study. *J. Cogn. Neurosci.* **26**, 1748–1763 (2014).
83. McGettigan, C. & Scott, S. K. Cortical asymmetries in speech perception: what’s wrong, what’s right and what’s left? *Trends Cogn. Sci.* **16**, 269–276 (2012).
84. McGettigan, C. *et al.* An application of univariate and multivariate approaches in fMRI to quantifying the hemispheric lateralization of acoustic and linguistic processes. *J. Cogn. Neurosci.* **24**, 636–652 (2012).
85. Recanzone, G. H. & Sutter, M. L. The Biological Basis of Audition. *Annu. Rev. Psychol.* **59**, 119–142 (2008).
86. Lewis, J. W. & Van Essen, D. C. Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J. Comp. Neurol.* **428**, 112–137 (2000).
87. Rauschecker, J. P. Cortical Processing Of Auditory Space: Pathways And Plasticity. in *Spatial Processing in Navigation, Imagery and Perception* (eds. Mast, F. & Jäncke, L.) 389–410 (Springer US, 2007). doi:10.1007/978-0-387-71978-8_20.

88. Vates, G. E. & Nottebohm, F. Feedback circuitry within a song-learning pathway. *Proc. Natl. Acad. Sci.* **92**, 5139–5143 (1995).
89. Friederici, A. Language Development and the Ontogeny of the Dorsal Pathway. *Front. Evol. Neurosci.* **4**, 3 (2012).
90. Jasmin, K., Lima, C. F. & Scott, S. K. Understanding rostral–caudal auditory cortex contributions to auditory perception. *Nat. Rev. Neurosci.* **20**, 425–434 (2019).
91. Hickok, G., Buchsbaum, B., Humphries, C. & Muftuler, T. Auditory–Motor Interaction Revealed by fMRI: Speech, Music, and Working Memory in Area Spt. *J. Cogn. Neurosci.* **15**, 673–682 (2003).
92. Pearce, M. T. & Wiggins, G. A. Expectation in Melody: The Influence of Context and Learning. *Music Percept.* **23**, 377–405 (2006).
93. Von Hippel, P. & Huron, D. Why Do Skips Precede Reversals? The Effect of Tessitura on Melodic Structure. *Music Percept. Interdiscip. J.* **18**, 59–85 (2000).
94. Russo, F. A. & Cuddy, L. L. A common origin for vocal accuracy and melodic expectancy: Vocal constraints. *J. Acoust. Soc. Am.* **105**, 1217–1217 (1999).
95. Schellenberg, E. G. Simplifying the Implication-Realization Model of Melodic Expectancy. *Music Percept. Interdiscip. J.* **14**, 295–318 (1997).
96. Zatorre, R. J. Pitch perception of complex tones and human temporal-lobe function. *J. Acoust. Soc. Am.* **84**, 566–572 (1988).
97. Johnsrude, I. S., Penhune, V. B. & Zatorre, R. J. Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* **123**, 155–163 (2000).

98. Meyer, M., Alter, K., Friederici, A. D., Lohmann, G. & von Cramon, D. Y. fMRI reveals brain regions mediating slow prosodic modulations in spoken sentences. *Hum. Brain Mapp.* **17**, 73–88 (2002).
99. Meyer, M., Steinhauer, K., Alter, K., Friederici, A. D. & von Cramon, D. Y. Brain activity varies with modulation of dynamic pitch variance in sentence melody. *Brain Lang.* **89**, 277–289 (2004).
100. Zatorre, R. J., Evans, A. C. & Meyer, E. Neural mechanisms underlying melodic perception and memory for pitch. *J. Neurosci.* **14**, 1908–1919 (1994).
101. Sammler, D., Grosbras, M.-H., Anwander, A., Bestelmeyer, P. E. G. & Belin, P. Dorsal and Ventral Pathways for Prosody. *Curr. Biol.* **25**, 3079–3085 (2015).
102. Jeffries, K. J., Fritz, J. B. & Braun, A. R. Words in melody: An H₂¹⁵O PET study of brain activation during singing and speaking. *NeuroReport Rapid Commun. Neurosci. Res.* **14**, 749–754 (2003).
103. Tervaniemi, M. & Hugdahl, K. Lateralization of auditory-cortex functions. *Brain Res. Rev.* **43**, 231–246 (2003).
104. Merrill, J. *et al.* Perception of Words and Pitch Patterns in Song and Speech. *Front. Psychol.* **3**, 76 (2012).
105. Sammler, D. & Elmer, S. Advances in the Neurocognition of Music and Language. *Brain Sci.* **10**, 509 (2020).
106. Albouy, P., Benjamin, L., Morillon, B. & Zatorre, R. J. Distinct sensitivity to spectrotemporal modulation supports brain asymmetry for speech and melody. *Science* **367**, 1043–1047 (2020).

107. Geiser, E., Zaehle, T., Jancke, L. & Meyer, M. The Neural Correlate of Speech Rhythm as Evidenced by Metrical Speech Processing. *J. Cogn. Neurosci.* **20**, 541–552 (2008).
108. Kasdan, A. V. *et al.* Identifying a brain network for musical rhythm: A functional neuroimaging meta-analysis and systematic review. *Neurosci. Biobehav. Rev.* **136**, 104588 (2022).
109. Grahn, J. A. & Brett, M. Rhythm and Beat Perception in Motor Areas of the Brain. *J. Cogn. Neurosci.* **19**, 893–906 (2007).
110. Rauschecker, J. P., Tian, B. & Hauser, M. Processing of Complex Sounds in the Macaque Nonprimary Auditory Cortex. *Science* **268**, 111–114 (1995).
111. Petkov, C. I. *et al.* A voice region in the monkey brain. *Nat. Neurosci.* **11**, 367–374 (2008).
112. Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P. & Pike, B. Voice-selective areas in human auditory cortex. *Nature* **403**, 309–312 (2000).
113. Pernet, C. R. *et al.* The human voice areas: Spatial organization and inter-individual variability in temporal and extra-temporal cortices. *NeuroImage* **119**, 164–174 (2015).
114. Agus, T. R., Paquette, S., Suied, C., Pressnitzer, D. & Belin, P. Voice selectivity in the temporal voice area despite matched low-level acoustic cues. *Sci. Rep.* **7**, 11526 (2017).
115. Latinus, M., Crabbe, F. & Belin, P. Learning-induced changes in the cerebral processing of voice identity. *Cereb. Cortex N. Y. N 1991* **21**, 2820–2828 (2011).
116. Ethofer, T. *et al.* Emotional voice areas: anatomic location, functional properties, and structural connections revealed by combined fMRI/DTI. *Cereb. Cortex N. Y. N 1991* **22**, 191–200 (2012).

117. Frühholz, S. & Grandjean, D. Processing of emotional vocalizations in bilateral inferior frontal cortex. *Neurosci. Biobehav. Rev.* **37**, 2847–2855 (2013).
118. Grossmann, T. The development of emotion perception in face and voice during infancy. *Restor. Neurol. Neurosci.* **28**, 219–236 (2010).
119. Angulo-Perkins, A. *et al.* Music listening engages specific cortical regions within the temporal lobes: Differences between musicians and non-musicians. *Cortex* **59**, 126–137 (2014).
120. Fedorenko, E., McDermott, J. H., Norman-Haignere, S. & Kanwisher, N. Sensitivity to musical structure in the human brain. *J. Neurophysiol.* **108**, 3289–3300 (2012).
121. Norman-Haignere, S., Kanwisher, N. G. & McDermott, J. H. Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron* **88**, 1281–1296 (2015).
122. Hickok, G. & Poeppel, D. The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).
123. Norman-Haignere, S. V. & McDermott, J. H. Neural responses to natural and model-matched stimuli reveal distinct computations in primary and nonprimary auditory cortex. *PLOS Biol.* **16**, e2005127 (2018).
124. Norman-Haignere, S., Kanwisher, N., McDermott, J. H. & Conway, B. R. Divergence in the Functional Organization of Human and Macaque Auditory Cortex Revealed by fMRI Responses to Harmonic Tones. *Nat. Neurosci.* **22**, 1057–1060 (2019).

125. Vanden Bosch der Nederlanden, C. M., Hannon, E. E. & Snyder, J. S. Finding the music of speech: Musical knowledge influences pitch processing in speech. *Cognition* **143**, 135–140 (2015).
126. Shoffstall, A. & Capadona, J. R. Chapter 28 - Prospects for a Robust Cortical Recording Interface. in *Neuromodulation (Second Edition)* (eds. Krames, E. S., Peckham, P. H. & Rezai, A. R.) 393–413 (Academic Press, 2018). doi:10.1016/B978-0-12-805353-9.00028-0.
127. Fifer, W. P. & Moon, C. M. The role of mother's voice in the organization of brain function in the newborn. *Acta Paediatr. Oslo Nor. 1992 Suppl.* **397**, 86–93 (1994).
128. Giordano, V. *et al.* Accent discrimination abilities during the first days of life: An fNIRS study. *Brain Lang.* **223**, 105039 (2021).
129. Mandel, D. R., Jusczyk, P. W. & Kemler Nelson, D. G. Does sentential prosody help infants organize and remember speech information? *Cognition* **53**, 155–180 (1994).
130. Kosakowski, H. *et al.* Preliminary evidence for selective cortical responses to music in one-month-old infants. Preprint at <https://doi.org/10.31219/osf.io/dq986> (2023).
131. Karmiloff-Smith, B. A. Beyond Modularity: A Developmental Perspective on Cognitive Science. *Eur. J. Disord. Commun.* **29**, 95–105 (1994).
132. Patel, A. Musicality and gene-culture coevolution: ten concepts to guide productive exploration. Preprint at <https://doi.org/10.31234/osf.io/qp6jx> (2021).
133. Kraus, N. & Banai, K. Auditory-Processing Malleability: Focus on Language and Music. *Curr. Dir. Psychol. Sci.* **16**, 105–110 (2007).
134. Vanden Bosch der Nederlanden, C. M. *et al.* Developmental changes in the categorization of speech and song. *Dev. Sci.* **26**, e13346 (2023).

135. Dehaene, S. Inside the Letterbox: How Literacy Transforms the Human Brain. *Cerebrum Dana Forum Brain Sci.* **2013**, 7 (2013).
136. Iuzzini-Seigel, J., Hogan, T. P., Rong, P. & Green, J. R. Longitudinal development of speech motor control: Motor and linguistic factors. *J. Mot. Learn. Dev.* **3**, 53–68 (2015).
137. Alcock, K. The development of oral motor control and language. *Syndr. Res. Pract. J. Sarah Duffen Cent.* **11**, 1–8 (2006).
138. Norman-Haignere, S. V. *et al.* Multiscale temporal integration organizes hierarchical computation in human auditory cortex. *Nat. Hum. Behav.* **6**, 455–469 (2022).
139. Patel, A. D. & Rueden, C. von. Where they sing solo: Accounting for cross-cultural variation in collective music-making in theories of music evolution. *Behav. Brain Sci.* **44**, e85 (2021).
140. Bruckert, L. *et al.* Vocal Attractiveness Increases by Averaging. *Curr. Biol.* **20**, 116–120 (2010).
141. Schneider, F. *et al.* Neuronal figure-ground responses in primate primary auditory cortex. *Cell Rep.* **35**, 109242 (2021).
142. Proctor, D. F. Modifications of Breathing for Phonation. in *Comprehensive Physiology* 597–604 (John Wiley & Sons, Ltd, 2011). doi:10.1002/cphy.cp030333.
143. Kayes, G. Structure and Function of the Singing Voice. in *The Oxford Handbook of Singing* (eds. Welch, G. F., Howard, D. M. & Nix, J.) 0 (Oxford University Press, 2019). doi:10.1093/oxfordhb/9780199660773.013.019.
144. Christiner, M. & Reiterer, S. Song and speech: examining the link between singing talent and speech imitation ability. *Front. Psychol.* **4**, 874 (2013).

145. Pfordresher, P. Q., Mantell, J. T. & Pruitt, T. A. Effects of intention in the imitation of sung and spoken pitch. *Psychol. Res.* **86**, 792–807 (2022).
146. Sundberg, J. Formant Structure and Articulation of Spoken and Sung Vowels. *Folia Phoniatr. Logop.* **22**, 28–48 (2009).
147. Rossi, S. *et al.* How the Brain Understands Spoken and Sung Sentences. *Brain Sci.* **10**, 36 (2020).
148. Leanderson, R., Sundberg, J. & Von Euler, C. Breathing muscle activity and subglottal pressure dynamics in singing and speech. *J. Voice* **1**, 258–261 (1987).
149. Salomoni, S., van den Hoorn, W. & Hodges, P. Breathing and Singing: Objective Characterization of Breathing Patterns in Classical Singers. *PLoS One* **11**, e0155084 (2016).
150. Hoit, J. D., Jenks, C. L., Watson, P. J. & Cleveland, T. F. Respiratory function during speaking and singing in professional country singers. *J. Voice* **10**, 39–49 (1996).
151. Nishimura, T. The descended larynx and the descending larynx. *Anthropol. Sci.* **126**, 3–8 (2018).
152. Bosma, J. F. *Symposium on Development of the Basicranium*. (U.S. Department of Health, Education, and Welfare, Public Health Service, National Institutes of Health, 1976).
153. Fowler, C. A. & Brown, J. M. Intrinsic f0 differences in spoken and sung vowels and their perception by listeners. *Percept. Psychophys.* **59**, 729–738 (1997).
154. Belyk, M. & Brown, S. Perception of affective and linguistic prosody: an ALE meta-analysis of neuroimaging studies. *Soc. Cogn. Affect. Neurosci.* **9**, 1395–1403 (2014).
155. Dichter, B. K., Breshears, J. D., Leonard, M. K. & Chang, E. F. The control of vocal pitch in human laryngeal motor cortex. *Cell* **174**, 21–31.e9 (2018).

156. Jürgens, U. The neural control of vocalization in mammals: a review. *J. Voice Off. J. Voice Found.* **23**, 1–10 (2009).
157. Callan, D. E. *et al.* Song and speech: Brain regions involved with perception and covert production. *NeuroImage* **31**, 1327–1342 (2006).
158. Warren, J. E., Wise, R. J. S. & Warren, J. D. Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends Neurosci.* **28**, 636–643 (2005).
159. Rauschecker, J. P. Ventral and dorsal streams in the evolution of speech and language. *Front. Evol. Neurosci.* **4**, 7 (2012).
160. von Holst, E. & Mittelstaedt, H. Das Reafferenzprinzip. *Naturwissenschaften* **37**, 464–476 (1950).
161. Bizley, J. K. & Walker, K. M. M. Sensitivity and Selectivity of Neurons in Auditory Cortex to the Pitch, Timbre, and Location of Sounds. *The Neuroscientist* **16**, 453–469 (2010).
162. Theunissen, F. E. & Elie, J. E. Neural processing of natural sounds. *Nat. Rev. Neurosci.* **15**, 355–366 (2014).
163. Eggermont, J. Periodicity pitch. in *Auditory Temporal Processing and its Disorders* (ed. Eggermont, J. J.) 0 (Oxford University Press, 2015).
doi:10.1093/acprof:oso/9780198719090.003.0008.
164. Cartwright, J. H. E., González, D. L. & Piro, O. Pitch perception: A dynamical-systems perspective. *Proc. Natl. Acad. Sci.* **98**, 4855–4859 (2001).
165. McPherson, M. J. & McDermott, J. H. Diversity in pitch perception revealed by task dependence. *Nat. Hum. Behav.* **2**, 52–66 (2018).

166. Imaizumi, S. *et al.* Vocal identification of speaker and emotion activates different brain regions. *NeuroReport* **8**, 2809 (1997).
167. Schirmer, A. & Kotz, S. A. Beyond the right hemisphere: brain mechanisms mediating vocal emotional processing. *Trends Cogn. Sci.* **10**, 24–30 (2006).
168. Morett, L. M. & Chang, L.-Y. Emphasising sound and meaning: pitch gestures enhance Mandarin lexical tone acquisition. *Lang. Cogn. Neurosci.* **30**, 347–353 (2015).
169. Rosen, S. M., Fourcin, A. J. & Moore, B. C. J. Voice pitch as an aid to lipreading. *Nature* **291**, 150–152 (1981).
170. Moore, B. C. J. *Hearing*. (Academic Press, 1995).
171. Uddin, L. Q. Chapter 1 - What Is Salience? in *Salience Network of the Human Brain* (ed. Uddin, L. Q.) 1–4 (Academic Press, 2017). doi:10.1016/B978-0-12-804593-0.00001-1.
172. Warren, J. D. & Griffiths, T. D. Distinct Mechanisms for Processing Spatial Sequences and Pitch Sequences in the Human Auditory Brain. *J. Neurosci.* **23**, 5799–5804 (2003).
173. Whitfield, I. C. Auditory cortex and the pitch of complex tones. *J. Acoust. Soc. Am.* **67**, 644–647 (1980).
174. Kazui, S., Naritomi, H., Sawada, T., Inoue, N. & Okuda, J.-I. Subcortical auditory agnosia. *Brain Lang.* **38**, 476–487 (1990).
175. Tramo, M. J., Shah, G. D. & Braida, L. D. Functional role of auditory cortex in frequency processing and pitch perception. *J. Neurophysiol.* **87**, 122–139 (2002).
176. Sankaran, N., Thompson, W. F., Carlile, S. & Carlson, T. A. Decoding the dynamic representation of musical pitch from human brain activity. *Sci. Rep.* **8**, 839 (2018).

177. Penagos, H., Melcher, J. R. & Oxenham, A. J. A Neural Representation of Pitch Salience in Nonprimary Human Auditory Cortex Revealed with Functional Magnetic Resonance Imaging. *J. Neurosci.* **24**, 6810–6815 (2004).
178. Warren, J. D., Uppenkamp, S., Patterson, R. D. & Griffiths, T. D. Separating pitch chroma and pitch height in the human brain. *Proc. Natl. Acad. Sci.* **100**, 10038–10042 (2003).
179. Belin, P., Bestelmeyer, P. E. G., Latinus, M. & Watson, R. Understanding Voice Perception. *Br. J. Psychol.* **102**, 711–725 (2011).
180. Andics, A., McQueen, J. M. & Petersson, K. M. Mean-based neural coding of voices. *NeuroImage* **79**, 351–360 (2013).
181. Kanber, E., Lavan, N. & McGettigan, C. Highly accurate and robust identity perception from personally familiar voices. *J. Exp. Psychol. Gen.* **151**, 897–911 (2022).
182. Puts, D. A., Gaulin, S. J. C. & Verdolini, K. Dominance and the evolution of sexual dimorphism in human voice pitch. *Evol. Hum. Behav.* **27**, 283–296 (2006).
183. Honjo, I. & Isshiki, N. Laryngoscopic and Voice Characteristics of Aged Persons. *Arch. Otolaryngol.* **106**, 149–150 (1980).
184. Boulet, M. J. & Oddens, B. J. Female voice changes around and after the menopause — an initial investigation. *Maturitas* **23**, 15–21 (1996).
185. Abdelli-Beruh, N. B., Wolk, L. & Slavin, D. Prevalence of Vocal Fry in Young Adult Male American English Speakers. *J. Voice* **28**, 185–190 (2014).
186. Dodd, B., Holm, A., Zhu, H. & Crosbie, S. Phonological development: A normative study of British English-speaking children. *Clin. Linguist. Phon.* **17**, 617–43 (2004).

187. Elardo, R., Bradley, R. & Caldwell, B. M. A Longitudinal Study of the Relation of Infants' Home Environments to Language Development at Age Three. *Child Dev.* **48**, 595–603 (1977).
188. Bloom, L. *et al.* Structure and Variation in Child Language. *Monogr. Soc. Res. Child Dev.* **40**, 1–97 (1975).
189. Fitch, W. T. & Giedd, J. Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *J. Acoust. Soc. Am.* **106**, 1511–1522 (1999).
190. Latinus, M., McAleer, P., Bestelmeyer, P. E. G. & Belin, P. Norm-Based Coding of Voice Identity in Human Auditory Cortex. *Curr. Biol.* **23**, 1075–1080 (2013).
191. Yan, W.-J., Wu, Q., Liang, J., Chen, Y.-H. & Fu, X. How Fast are the Leaked Facial Expressions: The Duration of Micro-Expressions. *J. Nonverbal Behav.* **37**, 217–230 (2013).
192. Conde, T. *et al.* The time course of emotional authenticity detection in nonverbal vocalizations. *Cortex* **151**, 116–132 (2022).
193. Sauter, D. A., Eisner, F., Calder, A. J. & Scott, S. K. Perceptual cues in nonverbal vocal expressions of emotion. *Q. J. Exp. Psychol.* **2006** **63**, 2251–2272 (2010).
194. Buck, R., Losow, J. I., Murphy, M. M. & Costanzo, P. Social facilitation and inhibition of emotional expression and communication. *J. Pers. Soc. Psychol.* **63**, 962 (19930401).
195. Provine, R. R. & Fischer, K. R. Laughing, Smiling, and Talking: Relation to Sleeping and Social Context in Humans. *Ethology* **83**, 295–305 (1989).
196. Hawkins, S., Cross, I. & Ogden, R. Communicative interaction in spontaneous music and speech. in *Language, music and interaction*. (eds. Orwin, M., Howes, C. & Kempson, R.) 285–329 (College Publications, 2013).

197. Ogden, R. & Hawkins, S. Entrainment as a basis for co-ordinated actions in speech. in *International Congress of the Phonetic Sciences* (York, 2015).
198. Cross, I. Music and communication in music psychology. *Psychol. Music* **42**, 809–819 (2014).
199. Garrod, S. & Pickering, M. J. Joint action, interactive alignment, and dialog. *Top. Cogn. Sci.* **1**, 292–304 (2009).
200. Kaukoma, T., Peräkylä, A. & Ruusuvuori, J. How Listeners Use Facial Expression to Shift the Emotional Stance of the Speaker's Utterance. *Res. Lang. Soc. Interact.* **48**, 319–341 (2015).

Acknowledgements

The authors thank the past and present members of the ICN's Speech and Communication laboratory for wonderfully stimulating lab meetings and the four reviewers for their extremely productive feedback and suggestions. The first author further thanks UCL for its generous GRS/ORS funding opportunity and I. Cross for his past mentorship and very interesting email discussions.

Author contributions

S.K.S. and I.H. researched data for the article, wrote the article and reviewed and/or edited the manuscript before submission. A.G. assisted with making the figures, and E.C.N. assisted with revising the text. All authors contributed substantially to discussion of the content.

Competing interests statement

The authors declare no competing interests.

Fig. 1: Comparison of spectra and breathing for song and speech

A schematic illustrative representation of the qualitative differences in the respiratory control required for and spectral characteristics of song and speech that have been reported in the literature. The same person sang then spoke the lyrics of a song, while their breathing patterns, laryngeal activity and voice were recorded. The top panel shows the harmonic structure of song, with the fundamental frequency marked in red and additional harmonics depicted by grey lines. The next panel shows the breathing pattern for the song, overlaid with the oscillogram in light blue (intensity of the song). The next panel shows the harmonic structure of speech, with the fundamental marked in red additional harmonics depicted by grey lines. The next panel shows the breathing pattern for the speech, overlaid with the oscillogram in light blue (intensity of the speech). Singing requires highly controlled outward airflow^{142,143} (depicted by black line, the green arrows indicate inhalations) in order to produce phonated vocalizations with sustained pitch⁴⁶ (shown by red line in top panel, which depicts F0, the fundamental frequency). Outward airflow in speaking is also highly controlled and supports fine-tuned pitch variations supporting the production of prosodic cues^{101,142}. Pitch contours – which underlie speech melody and sung melody – are shown in red and are more discrete for song than for speech⁴⁶, and behavioral studies suggest that it is easier to pitch-match during song imitation than during speech imitation due to song's more stable fundamental frequency^{144,145}. In general, songs contain longer vowel durations than speech¹⁴⁶, and sung sentences are slower than spoken ones^{47,147}, as shown in the schematic. Rapid and precise variations in subglottal pressure contribute to the production of sustained fundamental frequency in song and to pitch variations in emotive speech, whereas subglottal pressure is lower and less variable during baseline breathing and monotonous nonsense speech¹⁴⁸. Coordinated pre-phonatory inward movements of the muscles in the rib cage and abdomen are generally used by trained singers to increase intra-abdominal pressure during singing¹⁴⁹, though not for all musical styles (e.g., country¹⁵⁰). Both song and speech utilize dynamic filtering (as shown by the complex high-frequency spectral dynamics), where closure of the airway via myriad tongue, lip, and mandible movements facilitates the production of consonants, in order to convey meaning⁵⁵.

Fig. 2: Physiological and neural underpinnings of human vocalizations

a | The schematic depicts a singer's vocal tract while she is singing the same note but using different vowels (left image: /i/ and right image: /a/). In order to generate song, air from the lungs is pushed via the diaphragm through the vocal folds in the larynx (1) to generate a source signal. The source signal is subsequently filtered by the supralaryngeal vocal tract, which comprises a vertical (2) and horizontal cavity (3) of roughly the same length and volume that are connected by a narrow channel called the isthmus^{151,152}. Note the similar position of the larynx in both images—due to the same musical pitch being produced—but the very different shape of the globular tongue (4), whose differing shape impacts how the supralaryngeal vocal tract filters the signal and influences physical properties of the emitted sound, resulting in perceptually discriminable vowel sounds¹⁵³.

b | The neural bases of speech and song production largely overlap and are both underpinned by the pictured volitional vocalization network⁵. Spatially proximal specialized regions in the primary motor cortex drive laryngeal, respiratory, and articulatory motor neurons in the nucleus ambiguus⁵⁸. The laryngeal motor cortices (LMCs) are critical for producing volitional phonated sounds and contain ventral and dorsal subcomponents. Both subdivisions are involved in larynx-respiration coordination⁴⁹. The dorsal LMC shows activation for vertical larynx movements in the absence of speech¹⁵⁴, and has populations that selectively encode produced pitch¹⁵⁵. The ventral LMC is hypothesized to be less relevant for speech, but its functional role remains unknown⁴⁹. Sensory information from auditory and somatosensory cortices provides feedback inputs to the cortico-basal ganglia-thalamo-cortical loop via the insula to refine subsequent motor commands^{60,156}. The anterior insula is responsible for integrating sensory information from the auditory and somatosensory cortices to facilitate the generation of refined motor and sensory targets for articulation. Subcortical structures including the dorsal brainstem, pallidum, and putamen have also been shown to be recruited during overt volitional vocalizations⁵. Song production is generally associated with relatively increased right hemispheric activity in auditory^{48,102,157}, insular^{70,102}, and frontal regions⁴⁸ in comparison to speech (marked in blue) and there is evidence that singing may rely on right-lateralized auditory processing: the right anterior insula has been found to be important for sensorimotor integration during singing^{60,67,70}, and higher activity in the superior temporal gyrus (STG; particularly in the right hemisphere), part of the auditory cortex, has been found for singing vs. speaking^{48,157}.

Fig. 3: Cortical auditory processing: the dual-stream model and song selectivity

a | The rostral and caudal auditory streams allow for processing of 'what'- and 'where/how'-related auditory information, respectively^{77,90}. The rostral stream (also known as the ventral stream) identifies auditory objects and involves the primary auditory cortex (1), the anterior superior temporal sulcus (2) and the inferior frontal cortex (3)⁷⁷. This stream is involved in sound recognition and allows for categorization of speech phonemes and musical chords⁸¹. The caudal stream (also known as the dorsal stream) is involved in sensorimotor integration and spatial processing, and includes the planum temporale (4) which represents templates of 'do-able' sounds¹⁵⁸ and

interfaces with the inferior parietal lobe (5) and premotor cortex (6). This pathway underpins vocal learning for speech and song, integrating sensory information with efference copies of motor actions to allow for iterative fine-tuning of vocalizations^{159,160}.

b| If a population of neurons is selective for something, then the activity of this population must be indicative of the thing for which it selects¹⁶¹: This principle seems to hold for neuronal populations reported to encode the sound category of song, as shown in the schematic representation of the results from REF¹³. A song-selective population's response should somehow be indicative for music with singing: in this case, stronger, supra-additive responses were observed for music with singing (orange) than for instrumental music (green) or speech (pink).

Box 1: Pitch: Not a physical property, but a computed perceptual feature

Although natural sounds have heterogeneous acoustic properties, they can be characterized by the shared structure underlying fluctuations in their physical characteristics — for example, variations in the height (pitch) of natural sounds over time follow a shared power law relationship¹⁶². Many natural sounds, particularly vocalizations, exhibit spectral pitch (i.e., resolved harmonics¹⁶³) and periodicity (i.e., regular amplitude modulations), which gives rise to pitch¹⁶⁴ and facilitates the transmission of meaning and sound identity¹⁶⁵. Pitch cues from the voice support speaker identification (for example, by providing indications of their physical attributes¹⁶⁶), emotion communication (via prosodic cues, for example¹⁶⁷), word meaning (particularly in tonal languages¹⁶⁸), and sentence meaning (by providing segmental and suprasegmental information¹⁶⁹).

Although pitch is a key feature of complex periodic sounds, it is a perceptual attribute that cannot be denoted by physical units¹⁷⁰. For pure tones, the frequency of a sound can be used to determine its pitch and sounds with more regular periodic waveforms have salient pitch (i.e., attract attention and are behaviorally relevant¹⁷¹)¹⁶¹. For complex sounds, two spectral properties indicate pitch: the fundamental frequency (that is, the sound's lowest harmonic) and its harmonic structure (that is, the relations between elements in its harmonic series)¹⁶¹. For sounds having continuous spectra (such as iterated ripple noise), pitch can also be conferred via temporal or spectral regularities¹⁷².

Frequency-tuning is present in the cochlea and throughout the ascending auditory pathway, but pitch is a higher-level perceptual attribute that relies on cortical auditory neurons^{173–175}. Whereas the pitch of pure tones (corresponding to its frequency) can be identified by the activation of tuned neurons along the tonotopic map, the pitch of complex sounds is more computationally demanding and does not correspond to clear-cut cortical activity. For instance, pitch perception is influenced by context and two identical tones may be perceptually distinct, depending on the tone they are preceded by¹⁷⁶. Pitch discrimination requires the lateral Heschl's gyrus, which is sensitive to sound periodicity and encodes pitch salience^{61,177}. Responses to changes in pitch chroma (the pitch's musical note) and pitch height (how high or low a pitch is) occur in regions posterior and anterior to the primary auditory cortex, respectively¹⁷⁸.

Box 2: Voices and faces: two sides of the same coin?

Voices and faces are extremely salient stimuli within their respective sensory domains. Both provide information necessary for identity processing and the communication of emotion. Vocal and facial information can be used to determine similar attributes (such as identity, age, and gender^{179–181}). For example, ageing confers different physiologically determined vocal characteristics in men and women: puberty elicits highly dimorphic vocal pitch in men and women¹⁸²; vocal fold atrophy or edema due to ageing results in increased fundamental frequencies in men and (while these are decreased in women)¹⁸³; and women post-menopause experience difficulties exerting voice emission and high register vocalizations¹⁸⁴. Voice production is also influenced by socio-linguistic factors such as socio-economic status and cultural norms^{185–188}, impacting transmission of identity-based attributes. Gender can be

identified by vocal utterances of pre-pubescent children, when physiological differences are minimal¹⁸⁹.

Psychophysical and neural mapping studies have provided evidence for similar norm-based coding mechanisms and a similar organization of 'voice space' and 'face space' in the brain^{140,180,190}. However, the physical mechanics underlying voice perception constrain the vocal sounds that can be perceived to those which can be produced, whereas faces can be continuously perceived and our perception of them is not constrained to stimuli resulting from specific facial actions.

Emotions can be relayed by both the face and the voice, but emotional information is continuously 'leaked' via micro-expressions¹⁹¹, whereas transmission of emotion via the voice can be halted by remaining silent. However, if one does choose to vocalize, acoustic features are extremely indicative of emotion in both speech and reactive emotional vocalizations^{118,192,193}. Emotional expression in both modalities is differentially affected by our social situation. Social facilitation and inhibition occur for transmission of facial expression in social contexts, and rely on the closeness of the relationships between nearby individuals¹⁹⁴. Spontaneous smiling, laughter, and speech occurs to a greater extent when we are with others than when we are alone¹⁹⁵. When we engage in spontaneous phatic vocalizations — in which emotional content is more important than semantic content — we rhythmically align our voices with interlocutor(s)^{196,197}. Similar coordinative processes underscore the alignment of musical behaviors in group interactions¹⁹⁸ and evidence for automatic links between facial expressions, movements, and speech also exists^{199,200}.

Figure 1

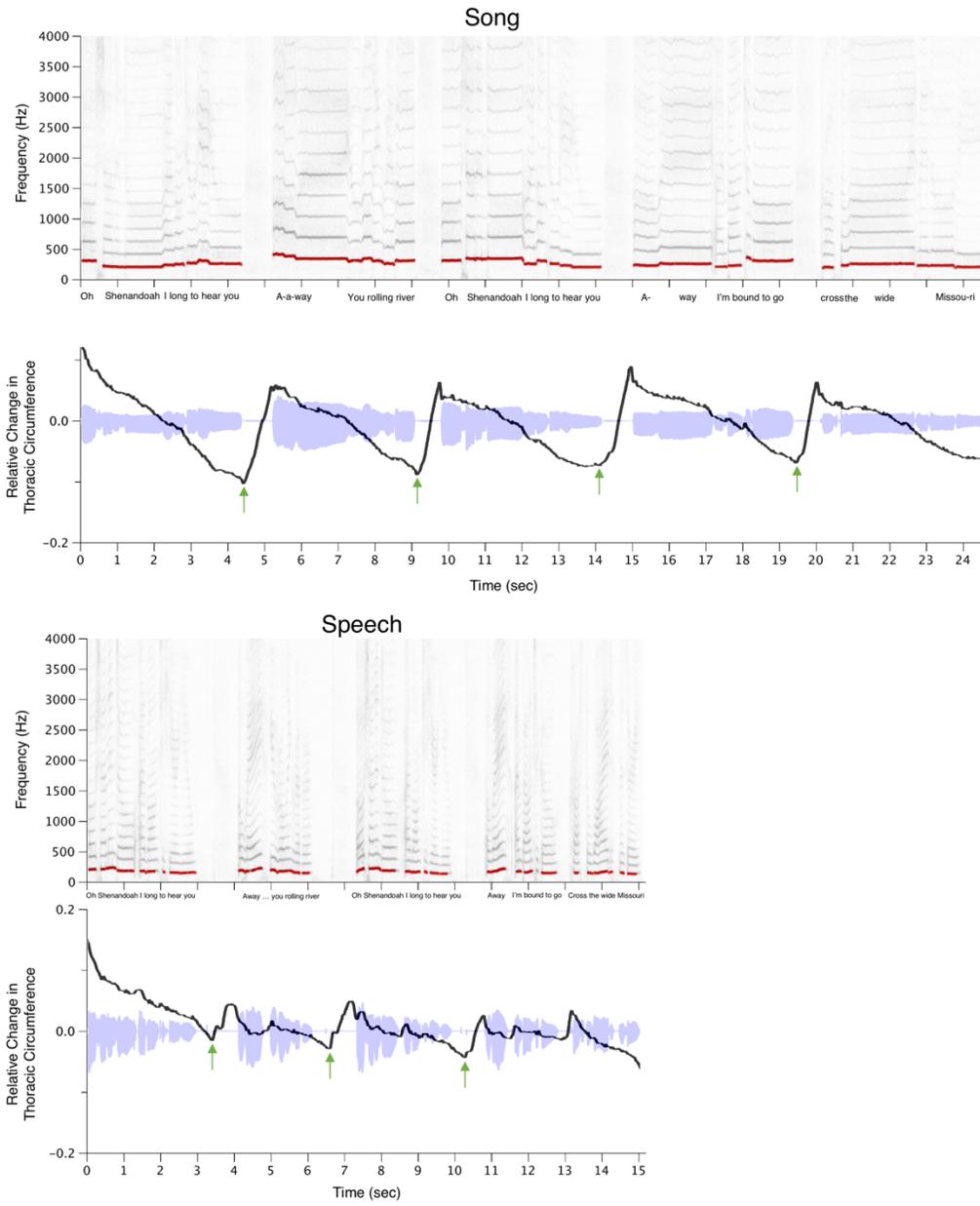
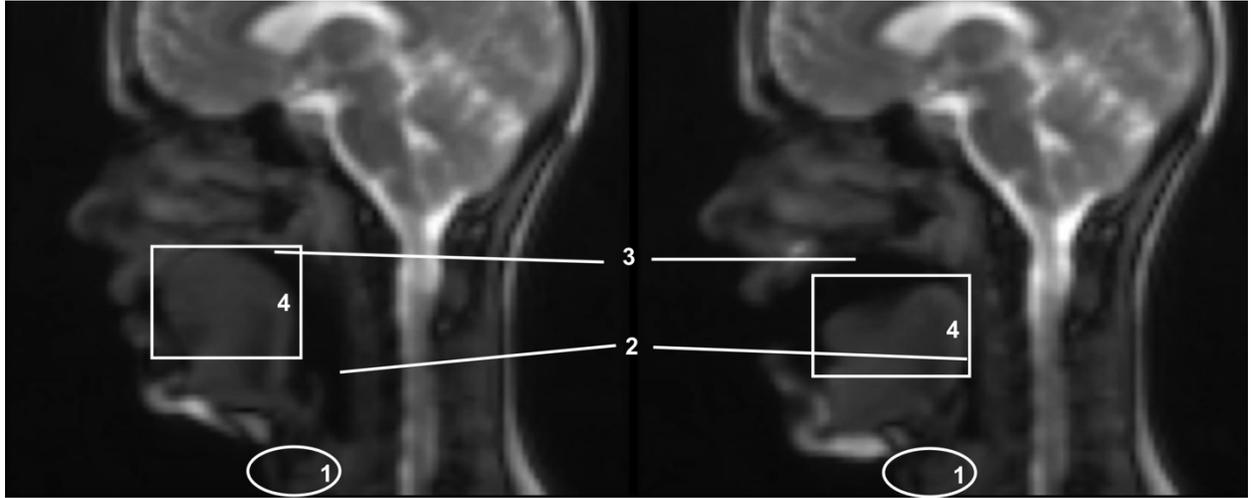
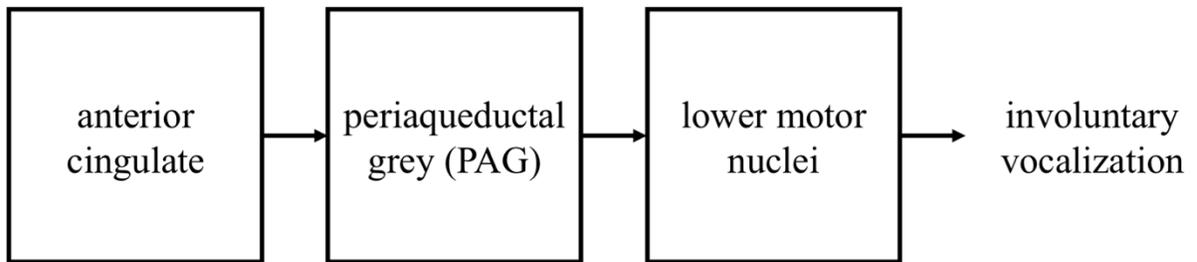


Figure 2
A



B



C

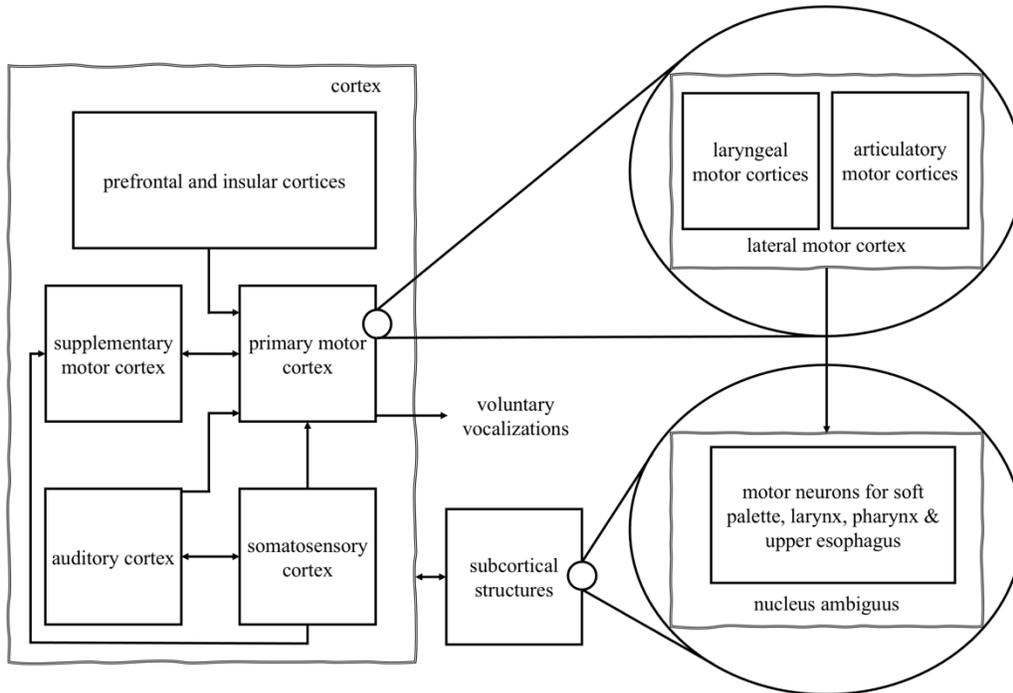


Figure 3

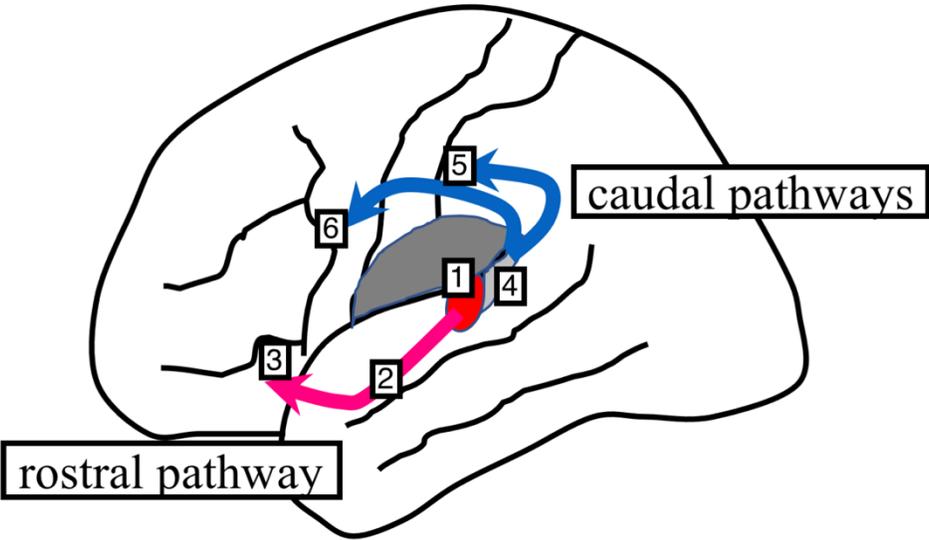
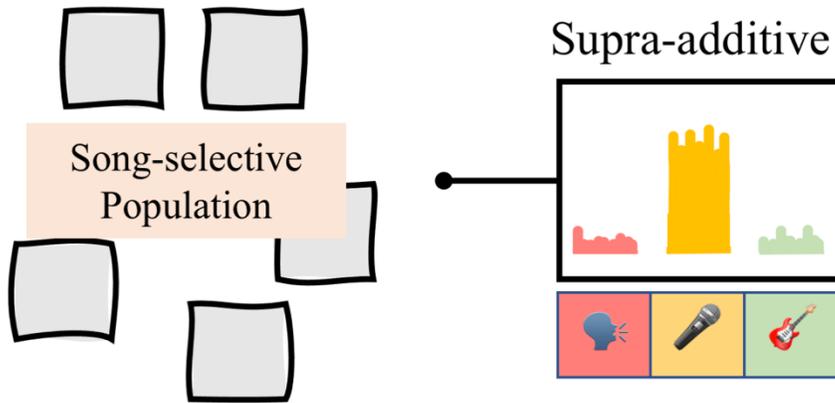


Figure 4

Indicative



Invariant

