*CORRESPONDENCE
Stefano Gonan
✉ stefano.gonan@phd.units.it
Giorgio Vallortigara
✉ giorgio.vallortigara@unitn.it
Cinzia Chiandetti
✉ cchiandetti@units.it

When sounds come alive: animacy in the auditory sense

Stefano Gonan [1]*, Giorgio Vallortigara [2]* and
Cinzia Chiandetti [1]*

[1]Department of Life Sciences, University of Trieste, Trieste, Italy, [2]CIMeC - Center for Mind/Brain
Sciences, University of Trento, Rovereto, Italy

Despite the interest in animacy perception, few studies have considered sensory modalities other than vision. However, even everyday experience suggests that the auditory sense can also contribute to the recognition of animate beings, for example through the identification of voice-like sounds or through the perception of sounds that are the by-products of locomotion. Here we review the studies that have investigated the responses of humans and other animals to different acoustic features that may indicate the presence of a living entity, with particular attention to the neurophysiological mechanisms underlying such perception. Specifically, we have identified three different auditory animacy cues in the existing literature, namely voicelikeness, consonance, and acoustic motion. While the first two characteristics are clearly exclusive to the auditory sense and indicate the presence of an animate being capable of producing vocalizations or harmonic sounds—with the adaptive value of consonance also being exploited in musical compositions in which the musician wants to convey certain meanings—acoustic movement is, on the other hand, closely linked to the perception of animacy in the visual sense, in particular to self-propelled and biological motion stimuli. The results presented here support the existence of a multifaceted auditory sense of animacy that is shared by different distantly related species and probably represents an innate predisposition, and also suggest that the mechanisms underlying the perception of living things may all be part of an integrated network involving different sensory modalities.

KEYWORDS

auditory perception, animacy, voicelikeness, acoustic motion, music

1 Introduction

Why are some sounds perceived as coming from living things and others not? Is there a sense of auditory animacy in different animal species? It is well known that animacy perception plays a crucial role in the survival of animals from an early age, enabling individuals to automatically and effortlessly locate biological cues of animate beings and then react according to the nature of the perceived entity, predator, prey, conspecific, etc. However, most research in this area has focused on the visual domain, particularly exploring the role of biological motion (Lorenzi et al., 2024), face perception (Kobylkov and Vallortigara, 2024; Kobylkov et al., 2024), and changes in speed and direction of moving objects (Di Giorgio et al., 2017; Lorenzi et al., 2017; review in Lorenzi and Vallortigara, 2021), while few researchers have addressed the possibility of an auditory counterpart to animacy perception (but see Tremoulet and Feldman, 2000).

This relative paucity of studies should not be surprising, however, given that auditory perception is characterized by a completely different psychophysical nature compared to the visual modality. As a result, the features that might constitute an animacy cue in the auditory sense are not necessarily the same as those that have been identified in visual research. When

we think about how animals, relying solely on their sense of hearing, are able to discriminate between living and non-living entities, the first feature that comes to mind is something unique to the auditory sense: the voice. Indeed, authors such as Patterson (2014) identify in the pulse-resonance structure of animal voices one of the key features at the basis of living organisms' recognition through sound.

It follows that all those sounds that are sufficiently similar in their acoustical structure to that of the voice—the so-called *voicelikeness*—should be able to elicit neurophysiological and behavioral responses similar to those of the voice (Broze, 2013). This was used by occultists at the turn of the 19th and 20th centuries, who argued for an ability allegedly demonstrated by certain people to perceive human voices in the noise produced by mechanical-electronic devices such as the radio, a phenomenon that became known as Electronic Voice Phenomena (Leary and Butler, 2015). They were actually documenting the auditory equivalent of pareidolia (a well-known phenomenon in vision), i.e., misperceiving or misinterpreting random noise or ambiguous acoustic stimuli as meaningful sounds, such as words or even sentences (Blom, 2015). This perceptual illusion suggests that the human species has a hypersensitivity to searching for and identifying voices within the auditory scene in which one is immersed, similar to what happens in the visual modality when faces are perceived in inanimate objects (Liu et al., 2014) also by non-human primates (Taubert et al., 2017). From a neuroanatomical perspective, the processing of voices of both conspecific and heterospecific individuals is supported by a dedicated extensive network of cortical areas (Broze, 2013).

In addition, another acoustic feature at the basis of the recognition of animate entities is consonance, i.e., sound frequencies that, when combined, produce the perception of a single sound and are therefore considered pleasant (Trulla et al., 2018). Because this feature is prominent in all acoustic events characterized by a well-defined and clear harmonic structure—such as the vocalizations of many species of songbirds and human speech—it is considered a hallmark of sounds produced by animate beings, which would explain the attraction that consonance exerts from an early age in distantly related taxa such as humans, domestic chicks and chimpanzees (Chiandetti and Vallortigara, 2011).

Finally, the perception of animacy in the auditory sense also occurs because of another acoustic feature that is closely related to one of the most important visual cues to animacy, i.e., locomotory movement. When animals move through space, they frequently simultaneously produce characteristic sounds—just think of the buzzing of a flying insect—defined as motion-induced sounds (Clark, 2016). These are involuntary acoustic events generated by the movement of anatomical parts of the body that are not specialized in producing communicative sounds, so that their conveyed information—i.e. the presence of a moving animal—is exploited by listeners rather than being transmitted directly and voluntarily (Clark, 2016). Neuroscientific studies have also demonstrated the existence of brain areas dedicated to the detection of acoustic motion, similar to those devoted to motion perception in other sensory modalities, such as vision (Wagner et al., 1997).

This work offers one of the first reviews in the field of auditory animacy perception and aims to fill this gap by systematizing the existing knowledge into a coherent description, identifying the main lines of research and proposing a global interpretation of the results obtained.

2 Voicelikeness

In vertebrates, voice production and perception are phenomena with a long evolutionary history. Voice generation mechanisms are the output of a system, the vocal tract, whose functional and neural control components are highly conserved across species (Newman, 2010; Fitch and Hauser, 2003). Furthermore, the ability to perceive conspecifics' voice is encoded in species-specific auditory nuclei in the auditory forebrain of birds (Louder et al., 2019) and, in the mammalian brain, regions such as the superior temporal plane and auditory neurons in the ventrolateral prefrontal cortex are involved in all studied primate species (Petkov et al., 2008; Romanski and Averbeck, 2009), with the anterior superior temporal sulcus specialized in humans (Belin et al., 2000; Fecteau et al., 2004). Moreover, the ability to recognize and approach the voice of a conspecific, especially if it is the mother's, is an ability shown early in life in many avian species, such as domestic chicks (Gottlieb, 1965; Fält, 1981; Bolhuis and van Kampen, 1991), pekin ducklings (Gottlieb, 1965), wood ducklings (Gottlieb, 1965), willow grouses (Allen, 1977), Japanese quails (Park and Balaban, 1991) and bobwhite quails (Barrow Heaton et al., 1978), which prefer the voice of their own species to that of another taxon or noise. As shown by Long et al. (2001), species-specific neural development underlies auditory preferences in taxa such as the domestic chicken and the Japanese quail: specifically, the authors transplanted developing neural tubes from embryonic quails to embryonic chickens and then tested the auditory preferences of the chimeric domestic chicks, finding that they began to prefer quail vocalizations.

However, voice-sensitive brain areas also recognize specific characteristics of the voice—such as fundamental frequency, call length and harmonic and phase-coupling content—regardless of the species. These neural correlates include the middle portions of the left and right superior temporal gyri, the right posterior superior temporal gyrus, the left Heschl's gyrus and left planum temporale in humans (Lewis et al., 2005; Bálint et al., 2023) and the mid and caudal ectosylvian gyri in dogs (Bálint et al., 2023). It can thus be posited that there exists an ancient neural predisposition to perceive sounds that exhibit characteristics that are analogous to those observed in acoustic events produced by vocal folds vibrations, a feature called *voicelikeness* (Schubert and Wolfe, 2016; Bálint et al., 2023). Consequently, auditory stimuli characterized by vocal similarity to certain innately encoded acoustic features will elicit a preference/approaching response. For example, one- and three-day-old domestic chicks run faster toward pure tones or tapping sounds when their frequency (Fischer, 1972), duration (Fischer, 1972), intensity (Fischer and Gilman, 1969), and rate (Fischer, 1972) are similar to those of the ideal maternal attraction call (Collias and Joos, 1953; Collias, 1987; Kent, 1993). In addition, sounds that emphasize a particular fundamental feature may be perceived as a superstimulus, and then be preferred over the original natural vocalization, as is the case with three-day-old chicks that run faster toward stimuli with a higher rate than natural mother's clucks (De Tommaso et al., 2019).

Most importantly, as demonstrated by Gilbert Gottlieb and colleagues in different auditory vs. visual choice experiments studying imprinting in domestic chicks (Gottlieb and Simner, 1969), pekin (Gottlieb and Klopfer, 1962; Klopfer and Gottlieb, 1962), mallard (Gottlieb, 1968) and wood ducklings (Gottlieb, 1968), the magnitude of attractiveness of vocal stimuli as animacy cues exceeds that of visual

stimuli, showing that—at least in the first days of life—the auditory recognition prevails, as ducklings prefer to follow a concealed moving loudspeaker broadcasting the species-specific maternal call instead of a silent moving visual replica of the hen, while domestic chicks are more attracted to an even simpler stimulus as an auditory flickering than to its visual version. Similarly, in the early stages of development, bobwhite quail chicks primarily rely on auditory cues for species identification and filial behavior, but as they mature, visual cues are integrated, with auditory stimuli remaining dominant (Lickliter and Virkar, 1989). Overall, these results support the cue hierarchy hypothesis proposed by Johnston and Gottlieb (1985), according to which sensory systems are organized hierarchically in early development, with the brain giving greater priority to auditory information as the auditory system matures faster than the visual system (Lickliter and Virkar, 1989).

3 Consonance

Analyses of the human voice have shown that when frequency and intensity values intersect, peaks corresponding to consonant melodic intervals emerge, suggesting that consonant intervals may represent the default state of human intonation (Schwartz et al., 2003). This phenomenon aligns with a distinctive characteristic of biological vocalizations: the presence of a well-defined harmonic structure, consisting of a fundamental frequency and harmonic overtones—a feature also shared by the sounds produced by musical instruments. This aspect was first noted in the 18th century by the French music theorist and composer Jean-Philippe Rameau, who observed that there was a connection between tonal sounds—often produced by living beings—and harmonic structure, particularly consonance (Christensen, 2004). Consonance is defined as the combination of two or more sound frequencies played simultaneously or consecutively that the brain perceives as stable, predictable and qualitatively pleasing. Conversely, dissonance results from the interaction of frequencies that when combined result as unstable, thus creating a perception of roughness or harshness in the auditory system.

This intuition has been confirmed by a number of studies of the vocalizations of different oscine species, such as the musician wren and the great tit, whose songs are characterized by a heavy use of consonant notes (Doolittle and Brumm, 2012; Richner, 2016), or, like the hermit thrush, which arranges its songs according to an overtone structure (Doolittle et al., 2014). Furthermore, in great tits, there is also a relationship between male fitness and the accuracy with which they produce consonant notes in their songs, showing that females prefer sounds characterized by stability and predictability (Richner, 2016). However, it is not even necessary for an animal to sing in order to "speak" consonant, since domestic chicks, for instance, emit perfect consonances across all types of calls, highlighting that consonant sounds are inherently present in animal communication (Maldarelli et al., 2024). Not only that, but the tone of our species' voice in dyadic interactions was shown preliminarily to be consonant when there is agreement between the speakers, and dissonant when disagreement arises (Okada et al., 2012).

Taken together, these results suggest a fundamental similarity between the harmonic structure of periodic sounds, whether the sound of the voice or that of musical instruments, thus making consonance a prominent aspect of vocalizations emitted by animate

beings (Bowling and Purves, 2015; Wagner and Hoeschele, 2022). This conclusion is further supported by the evidence that, in studies employing a spontaneous preference paradigm, consonant sounds are favored over dissonant ones in newborns of distantly related species, including domestic chickens (Chiandetti and Vallortigara, 2011), chimpanzees (Sugimoto et al., 2010) and humans (Masataka, 2006; Perani et al., 2010). This finding corroborates the idea that consonance is employed as a distinctive auditory cue for the presence of an animate object (Chiandetti and Vallortigara, 2011; Chiandetti, 2016; Vallortigara, 2021), as opposed to an inanimate one. Moreover, it suggests that this discriminative capacity may be a shared trait among species and may constitute an acoustic unlearned predisposition before experience, culture and training shape our preferences (Lahdelma and Eerola, 2020; Prete et al., 2020; Lahdelma et al., 2022).

4 Acoustic motion

The world is full of things that move and produce sound as a by-product. Some of them are alive—and the acoustic events produced are then called motion-induced sounds—and some of them are not, like leaves rustling, rocks tumbling or ocean waves. So, how does the auditory system discriminate between sounds produced by moving biological entities and those produced by moving objects? The first study to attempt to answer this question is that of Bidet-Caulet et al. (2005), in which the authors used fMRI technique to investigate in humans the neural correlates of auditory biological motion, more precisely the perception of footsteps. The results showed that the superior temporal sulcus was mainly involved in the processing of human motion sounds, irrespective of the sensory modality. A few years later, Cottrell and Campbell (2014) investigated auditory sensitivity to footsteps compared to non-biological impact sounds, such as a bouncing ball or drumbeats. Contrary to expectations, they found no increased sensitivity to biological motion compared to non-biological sounds, a finding that seems to indicate a difference between sensory modalities, with the visual domain being more fine-tuned when compared to the auditory domain. However, despite this discrepancy, both studies seem to converge on one aspect, namely the importance of temporal cues in the detection of human motion, especially if related to footsteps, a sound with high biological value (Bidet-Caulet et al., 2005; Cottrell and Campbell, 2014; Larsson, 2014).

The role of timing in the perception of auditory animacy was later taken up in the field of musical cognition, where there is a debate about the characteristics impacting the "aliveness" of a musical piece performance, an issue that, as some authors have noted (Blust et al., 2016), overlaps with that of auditory animacy. In particular, the bridge linking music to animacy is *rubato*, a performance technique used to convey greater expressiveness consisting of slight changes in the timing of a musical piece by slowing it down or speeding it up, giving also an acoustic motion sensation. In their 2015 paper, Blust and colleagues investigated the role of rubato in making music sound animated, asking participants to rate on a Likert scale the perceived animacy of computerized and human performances, both varying in rubato level. Results showed that both fixed and excessive timing variations led to a decrease in perceived animacy, while minimal levels of rubato resulted in significantly higher ratings of animacy (Blust et al., 2016).

The work of Nielsen et al. (2015), instead, made a significant contribution to the advancement of the field by investigating the effects of cues such as changes in the speed and direction of acoustic motion on the perception of animacy in humans. To this end, they designed a study that represents the auditory analog of the previously conducted research in the visual domain by Tremoulet and Feldman (2000). This was done by using binaural spatialisation to generate acoustic stimuli that mimicked the motion in three-dimensional space of a synthesized mosquito sound, varying in speed and direction—thus creating the impression of a living entity—or keeping the values constant in both dimensions. Participants were then asked to rate on a Likert scale their confidence that the perceived sound was lifelike or not, showing a significant difference between changes in speed and no changes, but unexpectedly also between changes in speed and changes in direction, suggesting the existence of a hierarchical organization of animacy cues with respect to acoustic motion. No study has yet addressed whether it is possible, in the acoustic domain, to distinguish causal interactions between inanimate and animate objects, as has been observed in the visual domain early in ontogenesis (Kominsky et al., 2022).

Finally, how do animals perceive acoustic motion? Even if animacy-related data are still absent, it is known that brain structures like the inferior colliculus are at the basis of the detection of acoustic motion direction, a process that is then refined in sensitivity through GABAergic inhibition, while motion information is then further processed by higher regions such as the auditory cortex (Wagner et al., 1997). Some animal taxa specialized in hunting using auditory cues even possess auditory space maps that integrate and help localize motion information (Wagner et al., 1997). (For a comprehensive review on this topic, see Carlile and Leung, 2016).

5 Discussion

This review has provided a full description of the area of auditory animacy, stressing the multifaceted nature of the acoustic mechanisms at the basis of animate beings' perception, going from voice features to sounds emitted involuntarily during biological motion. As outlined here, parallels between auditory and visual animacy cues are striking, with the voice being the acoustic analog of the face, and voicelikeness—resulting from the abstraction of vocal fundamental features—leading to a sensitivity comparable to that for face-like stimuli, causing the occurrence of phenomena such as auditory pareidolia, similar to face pareidolia. While the link between biological and acoustic movement is evident and very close, consonance as a feature of vocal prosody may instead fulfill a role akin to that of eyes and gaze, which are key characteristics in face and social perception.

Future research should consider whether face perception offers further analogies, such as in upside-down processing and backward speech, a domain where potential influences of rhythm, stress, and intonation can be examined (Toro et al., 2005). Studies should also address the potential effects of high-level cognition in modulating animacy perception—as done by Kim and Schachner (2021) who linked causal reasoning to animacy perception in music—and explore new potential cues, such as affective prosody (Zimmermann et al., 2013), or auditory regularities like fractal structure in vocal emissions (Jermyn et al., 2023). Finally, more research should investigate the cross modality of perceptual animacy, which may be part of a larger integrated system of social perception that includes different sensory modalities such as vision and hearing, and maybe even other unexplored domains, such as touch or olfaction.

Author contributions

SG: Writing – original draft, Conceptualization. GV: Writing – review & editing, Funding acquisition, Conceptualization. CC: Writing – review & editing, Funding acquisition, Conceptualization.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

Allen, H. M. (1977). The response of willow grouse chicks to auditory stimuli 1. Preference for hen grouse calls. *Behav. Process.* 2, 27–32. doi: 10.1016/0376-6357(77)90037-7

Bálint, A., Szabó, Á., Andics, A., and Gácsi, M. (2023). Dog and human neural sensitivity to voicelikeness: a comparative fMRI study. *Neuroimage* 265:119791. doi: 10.1016/j.neuroimage.2022.119791

Barrow Heaton, M., Miller, D. B., and Goodwin, D. G. (1978). Species-specific auditory discrimination in bobtail quail neonates. *Dev. Psychobiol.* 11, 13–21. doi: 10.1002/dev.420110106

Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., and Pike, B. (2000). Voice-selective areas in human auditory cortex. *Nature* 403, 309–312. doi: 10.1038/35002078

Bidet-Caulet, A., Voisin, J., Bertrand, O., and Fonlupt, P. (2005). Listening to a walking human activates the temporal biological motion area. *Neuroimage* 28, 132–139. doi: 10.1016/j.neuroimage.2005.06.018

Blom, J. D. (2015). Auditory hallucinations. *Handb. Clin. Neurol.* 129, 433–455. doi: 10.1016/B978-0-444-62630-1.00024-X

Blust, A. S., Baker, D. J., Richard, K., and Shanahan, D. (2016). Music, animacy, and rubato: what makes music sound human. In Proceedings of the 14th international Conference on Music Perception and Cognition. San Francisco, CA: ICMPC14, 186–190.

Bolhuis, J. J., and van Kampen, H. S. (1991). Auditory learning and filial imprinting in the chick. *Behaviour* 117, 303–319. doi: 10.1163/156853991X00607

Bowling, D. L., and Purves, D. (2015). A biological rationale for musical consonance. *Proc. Natl. Acad. Sci. USA* 112, 11155–11160. doi: 10.1073/pnas.1505768112

Broze, G. J. Y. III (2013). Animacy, anthropomimesis, and musical line. Columbus, OH: The Ohio State University.

Carlile, S., and Leung, J. (2016). The perception of auditory motion. *Trends in hearing.* 20:2331216516644254. doi: 10.1177/2331216516644254

Chiandetti, C. (2016). Commentary: cats prefer species-appropriate music. *Front. Psychol.* 7:594. doi: 10.3389/fpsyg.2016.00594

Chiandetti, C., and Vallortigara, G. (2011). Chicks like consonant music. *Psychol. Sci.* 22, 1270–1273. doi: 10.1177/0956797611418244

Christensen, T. (2004). Rameau and musical thought in the enlightenment. Cambridge: Cambridge University Press.

Clark, C. J. (2016). "Locomotion-induced sounds and sonations: mechanisms, communication function, and relationship with behavior" in Vertebrate sound production and acoustic communication. eds. R. Suthers, W. Fitch, R. Fay and A. Popper (Cham: Springer), 83–117.

Collias, N. E. (1987). The vocal repertoire of the red junglefowl: a spectrographic classification and the code of communication. *Condor* 89, 510–524. doi: 10.2307/1368641

Collias, N., and Joos, M. (1953). The spectrographic analysis of sound signals of the domestic fowl. *Behaviour* 5, 175–188. doi: 10.1163/156853953X00104

Cottrell, D., and Campbell, M. E. (2014). Auditory perception of a human walker. *Perception* 43, 1225–1238. doi: 10.1068/p7669

De Tommaso, M., Kaplan, G., Chiandetti, C., and Vallortigara, G. (2019). Naïve 3-day-old domestic chicks (*Gallus gallus*) are attracted to discrete acoustic patterns characterizing natural vocalizations. *J. Comp. Psychol.* 133, 118–131. doi: 10.1037/com0000132

Di Giorgio, E., Lunghi, M., Simion, F., and Vallortigara, G. (2017). Visual cues of motion that trigger animacy perception at birth: the case of self-propulsion. *Dev. Sci.* 20:12394. doi: 10.1111/desc.12394

Doolittle, E., and Brumm, H. (2012). O canto do Uirapuru: consonant intervals and patterns in the song of the musician wren. *J. Interdiscip. Music Stud.* 6, 55–85. doi: 10.4407/jims.2013.10.003

Doolittle, E. L., Gingras, B., Endres, D. M., and Fitch, W. T. (2014). Overtone-based pitch selection in hermit thrush song: unexpected convergence with scale construction in human music. *Proc. Natl. Acad. Sci. USA* 111, 16616–16621. doi: 10.1073/pnas.1406023111

Fält, B. (1981). Development of responsiveness to the individual maternal "clucking" by domestic chicks (*Gallus gallus domesticus*). *Behav. Process.* 6, 303–317. doi: 10.1016/0376-6357(81)90048-6

Fecteau, S., Armony, J. L., Joanette, Y., and Belin, P. (2004). Is voice processing species-specific in human auditory cortex? An fMRI study. *Neuroimage* 23, 840–848. doi: 10.1016/j.neuroimage.2004.09.019

Fischer, G. J. (1972). Sound stimuli and following in a domestic fowl: frequency, rate, and duration. *J. Comp. Physiol. Psychol.* 81, 183–190. doi: 10.1037/h0033533

Fischer, G. J., and Gilman, S. C. (1969). Following during imprinting as a function of auditory stimulus intensity. *Dev. Psychol.* 1, 216–218. doi: 10.1037/h0027303

Fitch, W. T., and Hauser, M. D. (2003). "Unpacking "honesty": vertebrate vocal production and the evolution of acoustic signals" in Acoustic communication. eds. S. AMED, R. R. Fay and A. N. Popper (New York: Springer-Verlag), 65–137.

Gottlieb, G. (1965). Imprinting in relation to parental and species identification by avian neonates. *J. Comp. Physiol. Psychol.* 59, 345–356. doi: 10.1037/h0022045

Gottlieb, G. (1968). Species recognition in ground-nesting and hole-nesting ducklings. *Ecology* 49, 87–95. doi: 10.2307/1933564

Gottlieb, G., and Klopfer, P. H. (1962). The relation of developmental age to auditory and visual imprinting. *J. Comp. Physiol. Psychol.* 55, 821–826. doi: 10.1037/h0041161

Gottlieb, G., and Simner, M. L. (1969). Auditory versus visual flicker in directing the approach response of domestic chicks. *J. Comp. Physiol. Psychol.* 67, 58–63. doi: 10.1037/h0026668

Jermyn, A. S., Stevenson, D. J., and Levitin, D. J. (2023). 1/f laws found in non-human music. *Sci. Rep.* 13:1324. doi: 10.1038/s41598-023-28444-z

Johnston, T. D., and Gottlieb, G. (1985). Development of visually controlled maternal preferences in Peking ducklings. *Dev. Psychobiol.* 18, 23–36. doi: 10.1002/dev.420180103

Kent, J. P. (1993). The chick's preference for certain features of the maternal cluck vocalization in the domestic fowl (*Gallus gallus*). *Behaviour* 125, 177–187. doi: 10.1163/156853993X00227

Kim, M., and Schachner, A. (2021). From music to animacy: causal reasoning links animate agents with musical sounds. In Proceedings of the Annual Meeting of the Cognitive Science Society. Austin, TX: Cognitive Science Society.

Klopfer, P. H., and Gottlieb, G. (1962). Imprinting and behavioral polymorphism: auditory and visual imprinting in domestic ducks (*Anas platyrhynchos*) and the involvement of the critical period. *J. Comp. Physiol. Psychol.* 55, 126–130. doi: 10.1037/h0045463

Kobylkov, D., Rosa-Salva, O., Zanon, M., and Vallortigara, G. (2024). Innate face-selectivity in the brain of young domestic chicks. *Proc. Natl. Acad. Sci. U. S. A.* 121:e2410404121. doi: 10.1073/pnas.2410404121

Kobylkov, D., and Vallortigara, G. (2024). Face detection mechanisms: Nature *vs.* nurture. *Front. Neurosci.* 18:1404174. doi: 10.3389/fnins.2024.1404174

Kominsky, J. F., Li, Y., and Carey, S. (2022). Infants' attributions of insides and animacy in causal interactions. *Cogn. Sci.* 46:e13087. doi: 10.1111/cogs.13087

Lahdelma, I., and Eerola, T. (2020). Cultural familiarity and musical expertise impact the pleasantness of consonance/dissonance but not its perceived tension. *Sci. Rep.* 10:8693. doi: 10.1038/s41598-020-65615-8

Lahdelma, I., Eerola, T., and Armitage, J. (2022). Is Harmonicity a misnomer for cultural familiarity in consonance preferences? *Front. Psychol.* 13:802385. doi: 10.3389/fpsyg.2022.802385

Larsson, M. (2014). Self-generated sounds of locomotion and ventilation and the evolution of human rhythmic abilities. *Anim. Cogn.* 17, 1–14. doi: 10.1007/s10071-013-0678-z

Leary, M. R., and Butler, T. (2015). "Electronic voice phenomena" in Parapsychology: a handbook for the 21st century. eds. E. Cardeña, J. Palmer and D. Marcusson-Clavertz (North Carolina: McFarland), 341–349.

Lewis, J. W., Brefczynski, J. A., Phinney, R. E., Janik, J. J., and DeYoe, E. A. (2005). Distinct cortical pathways for processing tool versus animal sounds. *J. Neurosci.* 25, 5148–5158. doi: 10.1523/JNEUROSCI.0419-05.2005

Lickliter, R., and Virkar, P. (1989). Intersensory functioning in bobwhite quail chicks: early sensory dominance. *Dev. Psychobiol.* 22, 651–667. doi: 10.1002/dev.420220702

Liu, J., Li, J., Feng, L., Li, L., Tian, J., and Lee, K. (2014). Seeing Jesus in toast: neural and behavioral correlates of face pareidolia. *Cortex* 53, 60–77. doi: 10.1016/j.cortex.2014.01.013

Long, K. D., Kennedy, G., and Balaban, E. (2001). Transferring an inborn auditory perceptual predisposition with interspecies brain transplants. *Proc. Natl. Acad. Sci. U. S. A.* 98, 5862–5867. doi: 10.1073/pnas.091080098

Lorenzi, E., Mayer, U., Rosa-Salva, O., and Vallortigara, G. (2017). Dynamic features of animate motion activate septal and preoptic areas in visually naïve chicks (*Gallus gallus*). *Neuroscience* 354, 54–68. doi: 10.1016/j.neuroscience.2017.04.022

Lorenzi, E., Nadalin, G., Morandi-Raikova, A., Mayer, U., and Vallortigara, G. (2024). Noncortical coding of biological motion in newborn chicks' brain. *Cereb. Cortex* 34:bhae262. doi: 10.1093/cercor/bhae262

Lorenzi, E., and Vallortigara, G. (2021). "Evolutionary and neural bases of the sense of animacy" in The Cambridge handbook of animal cognition. ed. J. C. Kaufman (New York: Cambridge University Press), 295–321.

Louder, M. I. M., Lawson, S., Lynch, K. S., Balakrishnan, C. N., and Hauber, M. E. (2019). Neural mechanisms of auditory species recognition in birds. *Biol. Rev. Camb. Philos. Soc.* 94, 1619–1635. doi: 10.1111/brv.12518

Maldarelli, G., Dissegna, A., Ravignani, A., and Chiandetti, C. (2024). Chicks produce consonant, sometimes jazzy, sounds. *Biol. Lett.* 20:20240374. doi: 10.1098/rsbl.2024.0374

Masataka, N. (2006). Preference for consonance over dissonance by hearing newborns of deaf parents and of hearing parents. *Dev. Sci.* 9, 46–50. doi: 10.1111/j.1467-7687.2005.00462.x

Newman, J. D. (2010). "Evolution of the communication brain in control of mammalian vocalization" in Handbook of mammalian vocalization - an integrative neuroscience approach. ed. M. B. Stefan (Amsterdam: Elsevier), 23–28.

Nielsen, R. H., Vuust, P., and Wallentin, M. (2015). Perception of Animacy from the motion of a single sound object. *Perception* 44, 183–197. doi: 10.1068/p7688

Okada, B. M., Lachs, L., and Boone, B. (2012). Interpreting tone of voice: musical pitch relationships convey agreement in dyadic conversation. *J. Acoust. Soc. Am.* 132:EL208–EL214. doi: 10.1121/1.4742316

Park, T., and Balaban, E. (1991). Relative salience of species maternal calls in neonatal gallinaceous birds: a direct comparison of Japanese quail (*Coturnix coturnix* japonica) and domestic chickens (*Gallus gallus* domesticus). *J. Comp. Psychol.* 105, 45–54. doi: 10.1037/0735-7036.105.1.45

Patterson, R. D. (2014). "Pulse-resonance sounds" in Encyclopedia of computational neuroscience. eds. D. Jaeger and R. Jung (New York, NY: Springer), 1–8.

Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., et al. (2010). Functional specializations for music processing in the human newborn brain. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4758–4763. doi: 10.1073/pnas.0909074107

Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., and Logothetis, N. K. (2008). A voice region in the monkey brain. *Nat. Neurosci.* 11, 367–374. doi: 10.1038/nn2043

Prete, G., Bondi, D., Verratti, V., Aloisi, A. M., Rai, P., and Tommasi, L. (2020). Universality vs experience: a cross-cultural pilot study on the consonance effect in music at different altitudes. *PeerJ* 8:e9344. doi: 10.7717/peerj.9344

Richner, H. (2016). Interval singing links to phenotypic quality in a songbird. *Proc. Natl. Acad. Sci. U. S. A.* 113, 12763–12767. doi: 10.1073/pnas.1610062113

Romanski, L. M., and Averbeck, B. B. (2009). The primate cortical auditory system and neural representation of conspecific vocalizations. *Annu. Rev. Neurosci.* 32, 315–346. doi: 10.1146/annurev.neuro.051508.135431

Schubert, E., and Wolfe, J. (2016). Voicelikeness of musical instruments: a literature review of acoustical, psychological and expressiveness perspectives. *Music. Sci.* 20, 248–262. doi: 10.1177/1029864916631393

Schwartz, D. A., Howe, C. Q., and Purves, D. (2003). The statistical structure of human speech sounds predicts musical universals. *J. Neurosci. Off. J. Soc. Neurosci.* 23, 7160–7168. doi: 10.1523/JNEUROSCI.23-18-07160.2003

Sugimoto, T., Kobayashi, H., Nobuyoshi, N., Kiriyama, Y., Takeshita, H., Nakamura, T., et al. (2010). Preference for consonant music over dissonant music by an infant chimpanzee. *Primates* 51, 7–12. doi: 10.1007/s10329-009-0160-3

Taubert, J., Wardle, S. G., Flessert, M., Leopold, D. A., and Ungerleider, L. G. (2017). Face pareidolia in the rhesus monkey. *Curr. Biol.* 27, 2505–2509.e2. doi: 10.1016/j.cub.2017.06.075

Toro, J. M., Trobalon, J. B., and Sebastián-Gallés, N. (2005). Effects of backward speech and speaker variability in language discrimination by rats. *J. Exp. Psychol. Anim. Behav. Process.* 31, 95–100. doi: 10.1037/0097-7403.31.1.95

Tremoulet, P. D., and Feldman, J. (2000). Perception of animacy from the motion of a single object. *Perception* 29, 943–951. doi: 10.1068/p3101

Trulla, L. L., Di Stefano, N., and Giuliani, A. (2018). Computational approach to musical consonance and dissonance. *Front. Psychol.* 9:381. doi: 10.3389/fpsyg.2018.00381

Vallortigara, G. (2021). Born knowing: imprinting and the origins of knowledge. Cambridge, MA: MIT Press.

Wagner, B., and Hoeschele, M. (2022). The links between pitch, timbre, musicality and social bonding from cross-species research. *Comp. Cogn. Behav.* 17, 13–32. doi: 10.3819/CCBR.2022.170002

Wagner, H., Kautz, D., and Poganiatz, I. (1997). Principles of acoustic motion detection in animals and man. *Trends Neurosci.* 20, 583–588. doi: 10.1016/s0166-2236(97)01110-7

Zimmermann, E., Leliveld, L., and Schehka, S. (2013). "Toward the evolutionary roots of affective prosody in human acoustic communication: a comparative approach to mammalian voices" in Evolution of emotional communication: from sounds in nonhuman mammals to speech and music in man, series in affective science. eds. E. Altenmüller, S. Schmidt and E. Zimmermann (Oxford: Oxford Academic).