

Article

Not peer-reviewed version

Human Sex Recognition: A Developmental Voice-Centered Model

[Yehuda Salu](#) *

Posted Date: 17 December 2025

doi: 10.20944/preprints202512.1610.v1

Keywords: human sex recognition; voice perception; developmental neuroscience; multisensory integration; auditory-visual learning; sexual development; hypothalamus; social cognition; biological sex cues; developmental models



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Human Sex Recognition: A Developmental Voice-Centered Model

Yehuda Salu

Howard University (Retired); yesalu@gmail.com

Abstract

Human sex recognition is a universal human ability, yet its biological and developmental basis remains incompletely characterized. Research on pheromonal signaling in animals has been essential for elucidating principles of sex recognition, social development, and multisensory integration across species, and has profoundly informed hypotheses about human behavior. However, in humans, no specific sex pheromone has been identified, and the anatomical and genetic substrates required for pheromonal communication—such as a functional vomeronasal organ and intact vomeronasal receptor gene families—are vestigial or absent. These observations suggest that additional or alternative mechanisms may play a central role in human sex recognition. This manuscript proposes the Human Sex Recognition (HSR) model, a developmental framework in which the human voice functions as the earliest and most reliable biological cue to sex. Drawing on converging observations reported across the scientific literature, the model integrates evidence that infants exhibit early sensitivity to vocal signals; that children form stable auditory–visual associations linking voices with sexed visual features; and that pubertal hypothalamic changes confer motivational and emotional significance on these learned associations. The manuscript is organized in two parts. Part I presents the formulation and developmental foundations of the HSR model. Part II evaluates empirical, biological, and comparative evidence bearing on human sex recognition and examines how the HSR framework complements existing pheromone-based sex recognition (HPSR) accounts. Although no new dataset is introduced, the synthesis of human developmental, perceptual, and neurobiological research indicates that HSR is supported by extensive real-life exposure, large effect sizes, and biologically plausible learning mechanisms. The HSR model incorporates mechanisms well established by pheromonal research while proposing additional developmental processes in domains where pheromonal explanations appear insufficient, thereby generating testable predictions for future experimental and developmental studies.

Keywords: human sex recognition; voice perception; developmental neuroscience; multisensory integration; auditory–visual learning; sexual development; hypothalamus; social cognition; biological sex cues; developmental models

Author Summary

Humans do not rely on pheromones to recognize the sex of others. Unlike animals that depend on chemical signals, humans develop sex recognition through hearing and vision, starting with the voice. Fetuses hear voices before birth, infants show early brain responses to voices, and children learn to match voice with visual cues. During puberty, sex hormones reorganize deep brain circuits that motivate sexual behavior, linking voice and vision to the hypothalamus. This developmental pathway: hearing and seeing → subcortical activation, forms a natural learning system for sex recognition. This article proposes a new model showing that the human voice provides the first and most reliable cue for identifying sex, offering a testable alternative to pheromone-based theories.

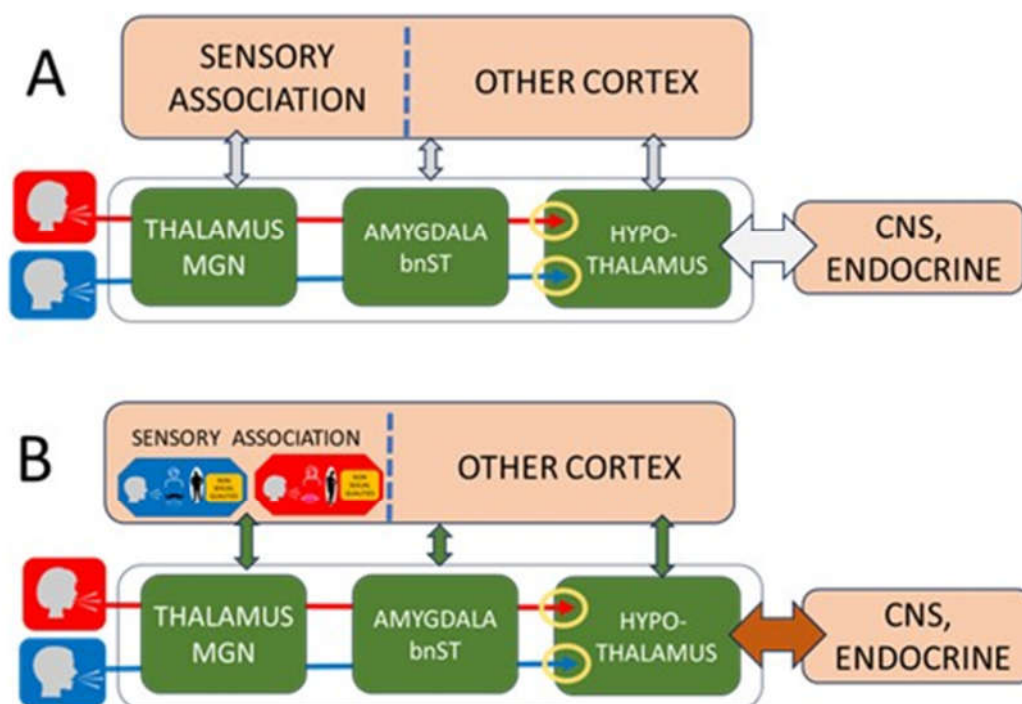


Figure 1. Visual Abstract.

Figure 1. Pheromone-based models propose that putative chemical cues are detected by the main olfactory epithelium and influence hypothalamic activity. However, the absence of a functional vomeronasal organ, pseudogenization of vomeronasal receptor genes, and the lack of any identified human sex pheromone reduce the biological plausibility of this pathway; reported effects on mate choice are weak and inconsistent. By contrast, the Human Sex Recognition (HSR) model emphasizes a developmental auditory–visual pathway.

A. Innate network. Innate auditory mechanisms distinguish male and female voices and relay this information through partially separate subcortical pathways that include the medial geniculate nucleus (MGN), amygdala, and bed nucleus of the stria terminalis (BNST), ultimately reaching hypothalamic regions that regulate sexual behavior (yellow circles). Neurons in these pathways connect with additional subcortical and cortical systems. Many of these connections are plastic and are modified by associative learning (vertical arrows).

B. Developmental integration. During childhood, associations form between auditory and visual cues corresponding to the speaker's sex. These cross-modal associations emerge first in sensory association cortex and later extend into broader cortical networks (green vertical arrows). During this period, both the hypothalamic–gonadal axis and the neural circuits linking sensory pathways to the developing sexual system are immature, and sexual behavior is not expressed. At puberty, gonadal hormones restructure connections between the hypothalamus, CNS, and reproductive organs (brown arrow), enabling sexual motivation and behavior to emerge.

Overview

Human sex recognition is an effortless and universal human ability, yet the biological and developmental mechanisms that support it remain incompletely characterized. Research on pheromonal signaling in rodents and other nonhuman mammals has been instrumental in establishing core principles of sex recognition, social communication, and neural specialization across

species, and has strongly influenced hypotheses about human sex recognition. On this basis, pheromonal cues have often been assumed to play a central role in humans, shaping interpretations of a limited set of behavioral and olfactory findings.

In humans, however, no specific sex pheromone has been identified, and the anatomical and genetic systems that mediate pheromonal communication in many nonhuman mammals—such as a functional vomeronasal organ and intact vomeronasal receptor families—are vestigial or absent. These human-specific biological constraints suggest that pheromonal mechanisms alone may be insufficient to account for the robustness, early emergence, and developmental continuity of human sex recognition, motivating the consideration of additional mechanisms grounded in human sensory access and neural development.

The **Human Sex Recognition (HSR) model** addresses this gap by proposing that the human voice serves as the earliest and most reliable biological cue to sex. In this account, infants exhibit early sensitivity to vocal signals, and children readily acquire stable auditory–visual associations linking sex-typical voices with corresponding visual features well before the onset of sexual behavior. At puberty, hormonal and neurodevelopmental changes reorganize hypothalamic and limbic pathways involved in motivation and affect, conferring new motivational significance on associations learned during earlier developmental stages. This developmental sequence—early auditory access, cross-modal learning in childhood, and pubertal neural remodeling—forms the core structure of the HSR model.

This manuscript has two aims. First, it presents the structure and developmental foundations of the HSR model and situates it within the broader context of human sensory, cognitive, and neural development. Second, it evaluates empirical and biological evidence relevant to human sex recognition and examines how the HSR framework complements pheromone-based sex recognition (HPSR) models by incorporating mechanisms well established in animal research while addressing domains in which pheromonal explanations are comparatively weak in humans. Although no new empirical data are introduced, the synthesis highlights systematic differences in the scope, effect sizes, and biological plausibility of the supporting evidence. Part I presents the theoretical and developmental formulation of the HSR model, and Part II examines comparative empirical findings and the biological constraints relevant to pheromone-based accounts.

Part I: The Human Sex Recognition (HSR) model

1. Introduction

Origins and Evolution of the Model

The Human Sex Recognition (HSR) model represents the most recent iteration of a theoretical framework developed over more than a decade to account for the biological and developmental foundations of human sex recognition. The earliest formulation (Salu, 2011) proposed that the human voice functions as a primary innate biological cue, serving as nature's agent for anchoring associations with other sensory inputs—particularly visual signals—during development. That work emphasized the role of prenatal auditory exposure and early postnatal learning in preparing later sexual behavior.

A subsequent refinement (Salu, 2020) focused on the contribution of subcortical structures to sexual development, highlighting the role of the medial geniculate nucleus (MGN), amygdala, bed nucleus of the stria terminalis (BNST), and hypothalamus in shaping sexual orientation and the expression of sexual motivation. This work emphasized the interaction between early sensory learning and later hormonal modulation of motivational circuits.

The present HSR framework integrates these earlier formulations within a unified developmental trajectory, informed by contemporary findings on cortical–subcortical interactions, multisensory learning, and neuroendocrine maturation. Rather than proposing a single moment of

sex recognition, the model describes a staged process in which early sensory access, experience-dependent learning, and pubertal neural reorganization interact to produce the robust and automatic capacity for sex recognition observed in humans.

AI-based language tools were used to assist with editorial clarity but did not contribute to the scientific content or theoretical development of the model.

2. Developmental Architecture of Human Sex Recognition

Prenatal foundations

The HSR model begins prenatally, when auditory exposure establishes the earliest biological substrate for sex recognition. By late gestation, the fetal auditory system is sufficiently mature to process external sounds, including speech. Fetuses demonstrate sensitivity to speech rhythms, discriminate sounds by pitch, and distinguish familiar from unfamiliar voices (Kisilevsky et al., 2003; Kisilevsky et al., 2009). Neurophysiological evidence confirms that the auditory pathway is functional during late gestation and responsive to external voices (Chelli & Chanoufi, 2008).

These findings suggest that the auditory system is primed before birth to encode sex-differentiated vocal features, such as pitch, timbre, and prosody. Importantly, this exposure occurs before the maturation of visual systems and long before the emergence of sexual behavior, positioning auditory cues as the earliest available scaffold for later sex recognition.

Auditory scaffold in infancy

Immediately after birth, infants exhibit specialized neural responses to voices. Neuroimaging studies show selective activation to vocal stimuli in early infancy (Grossmann et al., 2010), with evidence of voice-selective cortical processing by four months of age (Calce et al., 2024). By approximately 4–5 months, infants can discriminate between unfamiliar voices and reliably differentiate female voices (Yu et al., 2024).

These findings indicate that voice processing is not only present early but functionally specialized. Within the HSR model, the voice serves as the primary recognition axis during infancy, providing a stable and continuously available cue that can be accessed without visual attention, motor coordination, or explicit learning strategies.

Childhood cross-modal learning

As visual systems mature during infancy and early childhood, auditory information becomes increasingly integrated with visual cues. Cortical regions involved in face and biological motion processing—such as the fusiform face area (FFA), superior temporal sulcus (STS), and occipitotemporal cortex—support the binding of auditory and visual features into stable multimodal representations.

Empirical studies show that infants and young children can match voices to faces by sex (Richoz et al., 2017), and by the end of the first year of life, they integrate vocal cues with biological motion and body form (Johnson et al., 2021). In some contexts, vocal information may even outweigh visual cues for social judgments (Blasi et al., 2023). These findings support the view that sex recognition in childhood is not tied to a single modality but emerges from experience-dependent cross-modal learning organized around an early auditory scaffold.

Within the HSR framework, this period is critical for establishing durable associations between voices and sex-typical visual features. These associations are formed well before sexual motivation emerges and are shaped by extensive real-life exposure rather than explicit instruction.

Pubertal expansion and motivational reorganization

Puberty introduces a major transformation in the functional significance of previously learned sensory associations. Gonadal hormones reorganize hypothalamic and limbic circuits involved in

motivation, affect, and social behavior (Sisk & Zehr, 2005). In parallel, the human voice undergoes pronounced pubertal changes, increasing its salience as a sex-differentiated cue (Salu, 2011). These changes are universal and perceptually obvious, requiring no specialized instrumentation to detect.

Beyond these salient changes, acoustic analyses reveal that subtler sex-related differences in vocal parameters emerge gradually between approximately 5 and 15 years of age, allowing reliable classification of a child's sex from speech alone (Chen et al., 2022). Neuroimaging evidence further indicates that adolescents show increased hypothalamic responsiveness to visual sex cues compared to children, reflecting a developmental shift in motivational weighting across modalities (Nowling et al., 2024).

Within the HSR model, puberty does not initiate sex recognition but reorganizes its motivational significance. Associations established earlier through auditory–visual learning are now integrated into hypothalamic–limbic circuits, supporting sexual attraction, orientation, and behavior.

Part II: Evaluation, Complementarity, and Testable Implications

3. Complementary Evaluation of Human Sex Recognition Models

Research on pheromonal signaling has played a foundational role in elucidating sex recognition and social communication in nonhuman mammals, providing a rich empirical framework for understanding how biologically relevant cues can guide development and behavior. These findings have naturally informed hypotheses about human sex recognition, leading to models that emphasize chemosensory mechanisms as potential contributors to human sexual development.

In humans, several studies have reported hypothalamic activation in response to putative chemosignals such as androstadienone and estratetraenol (Savic et al., 2001; Savic & Lindström, 2008). These findings suggest that chemosensory inputs may influence neural systems involved in arousal, affect, or social evaluation. However, behavioral effects associated with these compounds are variable, context-dependent, and not consistently linked to explicit sex recognition. Moreover, no universal human sex pheromone has been identified, and the developmental role of chemosensory cues in early sex recognition remains unclear.

The HSR model addresses these limitations by focusing on sensory channels that are developmentally accessible from the prenatal period onward and that support continuous, high-fidelity learning through real-world exposure. Rather than rejecting pheromonal mechanisms, the HSR framework incorporates insights from animal research while proposing that, in humans, auditory and visual cues constitute the primary scaffolding for sex recognition, with chemosensory influences acting, at most, as modulators of affect or motivation.

4. Scientific Criteria and Model Evaluation

A useful scientific model must be biologically feasible, developmentally grounded, and empirically falsifiable. The HSR model satisfies these criteria by generating specific predictions tied to sensory access, neural development, and learning history.

Because the model is grounded in well-characterized developmental processes—prenatal auditory sensitivity, infant voice selectivity, cross-modal learning, and pubertal neuroendocrine reorganization—it yields predictions that can be tested using established behavioral, neuroimaging, and developmental methods. Importantly, the model predicts differential outcomes under sensory deprivation conditions, allowing it to be empirically distinguished from models that emphasize chemosensory primacy.

By contrast, pheromone-based models in humans face greater challenges in specifying developmental mechanisms, identifying invariant stimuli, and generating clear falsifiable predictions about early sex recognition. This difference does not diminish the importance of pheromonal research in animals, but it highlights the need for models that reflect human-specific sensory ecology and developmental trajectories.

5. Predictions and Testable Hypotheses

The HSR model generates a set of empirically testable predictions:

1. Developmental timing

Children should reliably classify sex by voice earlier than by face, reflecting earlier auditory access and learning.

2. Sensory deprivation

Congenital or early auditory deprivation should impair the development of sex recognition more strongly than olfactory impairment.

3. Pubertal reorganization

Adolescents should exhibit increased hypothalamic and limbic responsiveness to visual sex cues compared to children, reflecting the integration of earlier learned associations into motivational circuits.

4. Chemosensory predictions

If pheromone-based models are correct, olfactory deprivation should directly impair sex recognition from auditory or visual cues. Although anosmia and smell loss are associated with changes in sexual behavior, intimacy, or mood (Croy et al., 2013; Oleszkiewicz et al., 2020), no studies to date demonstrate a direct impairment in voice- or face-based sex recognition in individuals with olfactory deficits.

5. Neural pathway differentiation

The HSR model predicts that male and female voices may engage partially distinct neural pathways from the medial geniculate nucleus (MGN) to hypothalamic and limbic structures, with possible overlap. Existing neuroimaging studies already demonstrate sex-differentiated cortical processing of voices, including tonotopic organization (Talavage et al., 2004), hemispheric asymmetries (Lattner et al., 2004), sex-specific voice activations (Sokhi et al., 2005), gender-sensitive processing (Charest et al., 2012), and distinct representations of pitch and timbre (Allen et al., 2016). These approaches could be extended to test subcortical pathway differentiation predicted by the HSR framework.

6. Chemosensory Influences in Humans: A Modulatory Role

Some authors have argued that, even in the absence of a functional vomeronasal organ, humans may process socially relevant chemosignals via the main olfactory system (Herz & Inzlicht, 2014). Such signals may influence bonding, stress responses, or emotional context. Within the HSR framework, these effects are fully compatible with a modulatory role for chemosensory inputs.

However, as emphasized by Trotier (2011), human sexuality involves a complex integration of psychological, cognitive, and social processes that extend beyond the pheromonal mechanisms governing mate recognition in many nonhuman species. From this perspective, chemosensory cues may shape emotional tone or arousal but are unlikely to function as primary recognition mechanisms for sex in humans.

7. Biological Feasibility and Comparative Perspective

Biological feasibility is a critical criterion for evaluating any model of human behavior. The HSR model aligns with comparative evidence across vertebrates, where auditory signals—such as song in birds or calls in amphibians—often play a central role in sex recognition. These parallels suggest that reliance on acoustic cues for sex recognition is evolutionarily conserved and biologically parsimonious.

In humans, pubertal voice changes driven by sex hormones further underscore the biological preparation of vocal cues for recognizing sexually mature individuals. Visual cues may provide redundant or complementary information, enhancing robustness. Together, these features support the biological plausibility of the HSR model while remaining consistent with insights derived from animal research on communication and recognition.

8. Interpreting Effect Size, Sample Scale, and Error Bounds

The comparison summarized in Table X relies on effect size rather than statistical significance alone. In particular, the h statistic is used as a standardized measure of the magnitude of a difference between proportions, independent of sample size. This distinction is important because statistical significance can be achieved with small effects in small samples, while effect size reflects the practical and perceptual strength of a phenomenon.

In studies of putative human chemosignals, reported effects are typically modest and derived from relatively small, laboratory-based samples. As a result, confidence intervals remain wide, and replication across contexts has been inconsistent. By contrast, the Human Sex Recognition (HSR) model is supported by developmental phenomena that are expressed continuously across everyday human interaction, beginning in infancy and extending across the lifespan. These include voice-based sex classification, cross-modal voice–face matching, and pubertal vocal differentiation, all of which are observed repeatedly across cultures and contexts.

Although these observations are not derived from a single experimental dataset, their effective sample size is extraordinarily large, encompassing countless naturalistic exposures accumulated over development. From a statistical perspective, such massive observational scale substantially constrains plausible error ranges, even when individual experimental effect sizes are moderate. This distinction between limited laboratory sampling and pervasive real-world exposure is central to evaluating the relative robustness of competing models.

Accordingly, the comparison in Table X should not be interpreted as a contest between isolated experiments, but as a contrast between narrowly sampled effects and developmentally entrenched phenomena whose reliability is supported by scale, consistency, and biological plausibility.

9. Conclusion

Human sex recognition emerges from a developmental trajectory organized around an auditory–visual–subcortical axis. Prenatal and early postnatal exposure to voices provides the initial scaffold; childhood experience supports cross-modal learning; and pubertal neuroendocrine changes reorganize motivational circuits that confer sexual significance on earlier learned associations.

The HSR model integrates mechanisms corroborated by pheromonal research while proposing additional developmental processes that address domains in which pheromonal explanations are comparatively weak in humans. By generating falsifiable predictions and grounding sex recognition in human-specific sensory access and neural development, the HSR framework offers a complementary and biologically grounded account of the human sex instinct.

Table 1. Interpreting Effect Size, Observational Scale, and Ecological Context.

Dimension	HSR (Human Voice-Based Sex Recognition)	HPSR (Human Pheromone-Based Sex Recognition)
Typical experimental samples	50–500 participants per study; supplemented by continuous naturalistic exposure across development	20–60 participants per laboratory study
Effective observational scale	Extremely large (lifelong exposure across millions of real-world interactions)	Limited to episodic laboratory exposure
Typical classification accuracy	~90–98% for voice-based sex classification	~55–60%, often near chance
Cohen's h (effect size)	Large to very large ($h \approx 1.1$ – 1.3), reflecting robust sex classification	Very small ($h \approx 0.05$ – 0.25), reflecting subtle or context-dependent effects
Biological mechanisms engaged	Auditory, cortical, subcortical, and developmental learning systems	Chemosensory pathways with unclear developmental primacy in humans

Comparison of evidence supporting voice-based human sex recognition (HSR) and pheromone-based sex recognition (HPSR), considering effect size (h), experimental sample size, effective observational scale, ecological context, and biological plausibility. Effect size estimates reflect representative magnitudes reported in the literature for sex classification or sex-linked neural responses, rather than pooled meta-analytic estimates.

The comparison summarized in Table 1 emphasizes effect size and inferential robustness rather than statistical significance alone. Cohen's h is reported as a standardized measure of the magnitude of differences in classification performance, independent of sample size. This distinction is critical because statistically significant results may reflect small effects in limited samples, whereas effect size provides a more direct indication of perceptual and behavioral reliability.

In studies supporting pheromone-based accounts, effects are typically derived from laboratory paradigms involving instructed sniffing of isolated chemosensory samples in relatively small participant groups. These methods are well suited for detecting subtle physiological or affective modulation under controlled conditions but necessarily abstract chemosensory cues from the multisensory and social environments in which human sex recognition normally occurs. As a result, confidence intervals remain wide and replication across contexts has been variable.

By contrast, the HSR model is supported by phenomena that operate continuously in naturalistic settings, including voice-based sex classification, cross-modal voice–face matching, and developmentally embedded learning beginning in infancy. Although individual experimental studies may involve modest sample sizes, the effective observational scale is extremely large due to the accumulation of real-world exposure across development. From an inferential perspective, such

scale substantially constrains plausible error ranges, even when individual experimental effect sizes are moderate.

Accordingly, the comparison in Table X should be interpreted not as a contest between isolated experiments, but as a contrast between episodic laboratory effects and developmentally entrenched mechanisms whose reliability is supported by scale, consistency, and biological plausibility.

Note on effect size estimates.

Cohen's h values are reported as approximate ranges intended to convey the typical magnitude of differences observed in voice-based versus chemosensory paradigms relevant to human sex recognition. For HSR, h values in the large to very large range ($h \approx 1.1$ – 1.3) are consistent with reported classification accuracies of ~90–98% for voice-based sex identification across languages and contexts, relative to chance performance. For HPSR, h values in the very small range ($h \approx 0.05$ – 0.25) reflect modest deviations from chance commonly reported in laboratory studies examining behavioral or neural responses to isolated chemosensory compounds (e.g., hypothalamic activation or forced-choice judgments), rather than explicit sex classification performance. These ranges are illustrative of relative effect magnitude and robustness, not formal pooled estimates.

AI Assistance Disclosure

The author used AI-based language tools to assist with editing, restructuring, and clarity of presentation. All scientific content, theoretical framework, interpretations, and conclusions are the author's own and are based on prior published work and cited empirical literature.

References

1. Allen EJ, Burton PC, Olman CA and Oxenham AJ. (2016). Representations of pitch and timbre variation in human auditory cortex. *The Journal of Neuroscience* 37:1284–1293. doi: 10.1523/JNEUROSCI.2504-16.201
2. Belin, P., Zatorre, R., & Ahad, P. (2002). Human temporal-lobe regions sensitive to voice. *Nature Neuroscience*, 5, 965–970.
3. Blasi CH, et al. (2023). Young children's attributes are better conveyed by voices than by faces. *Journal of Experimental Child Psychology*. doi: 10.1016/j.jecp.2022.105530
4. Calce RP, et al. (2024). Voice categorization in the four-month-old human brain. *Current Biology* 34:292–299. doi: 10.1016/j.cub.2023.11.019
5. Campanella S and Belin P. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences* 11:535–543.
6. Charest I, Pernet C, Latinus M, Crabbe F and Belin P. (2012). Cerebral processing of voice gender studied using a continuous carryover fMRI design. *Cerebral Cortex* 23:958–966. doi: 10.1093/cercor/bhs090
7. Chelli D and Chanoufi B. (2008). Fetal audition: Myth or reality? *Journal de Gynécologie Obstétrique et Biologie de la Reproduction* 37:554–558. doi: 10.1016/j.jgyn.2008.06.007
8. Chen L, Togneri R, Maybery M and Tan T. (2022). Automated sex classification of children's voices and changes in differentiating factors with age. arXiv preprint. Available at: <https://arxiv.org/abs/2209.13112>
9. Croy I, Bojanowski V and Hummel T. (2013). Men without a sense of smell exhibit a strongly reduced number of sexual relationships, women exhibit reduced partnership security – a reanalysis of previously published data. *Biological Psychology* 92:292–294. doi: 10.1016/j.biopsycho.2012.09.008
10. DeCasper, A. J., & Fifer, W. (1980). Of human bonding: Newborns prefer their mother's voice. *Science*, 208, 1174–1176.
11. DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborn listening preferences. *Infant Behavior and Development*, 9, 133–150.

12. Dulac C and Torello AT. (2003). Molecular detection of pheromone signals in mammals. *Nature Reviews Neuroscience* 4:551–562.
13. Gluckman M and Johnson S. (2013). Early developmental differences in voice and face recognition. *Developmental Science* 16:739–748.
14. Grossmann T, Oberecker R, Koch SP and Friederici AD. (2010). The developmental origins of voice processing in the human brain. *Neuron* 65:852–858.
15. Herz RS and Inzlicht M. (2014). Sex differences in response to physical and social cues: A review. *Hormones and Behavior* 66:545–560. doi: 10.1016/j.yhbeh.2014.10.001
16. Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099–3111
17. Hummer, T. A., & McClintock, M. K. (2009). Putative human pheromones and their effects on attention. *Psychoneuroendocrinology*, 34, 142–153.
18. Jacob S, et al. (2002). Androstadienone modulates women’s mood and cortisol. *Journal of Neuroscience* 22:239–245.
19. Johnson SP, Dong M, Ogren M and Senturk D. (2021). Infants’ identification of gender in biological motion displays. *Infancy* 26:687–706.
20. Kisilevsky BS, et al. (2003). Effects of experience on fetal voice recognition. *Psychological Science* 14:220–224.
21. Kisilevsky B, et al. (2009). Fetal sensitivity to properties of maternal speech and language. *Infant Behavior and Development* 32:59–71.
22. Lattner S, Meyer ME and Friederici AD. (2004). Voice perception: Sex, pitch, and the right h Lecanuet, J. P., & Schaal, B. (1996). Fetal sensory competencies. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 68, 1–8.emisphere. *Human Brain Mapping* 24:11–20. doi: 10.1002/hbm.20065
23. Meredith M. (2001). Human vomeronasal organ: Vestigial or functional? *Chemical Senses* 26:433–445.
24. Nowling D, et al. (2024). Sex differences in development of functional connections in the face processing network. *NeuroImage* 283:120384.
25. Oleszkiewicz A, et al. (2020). Consequences of undetected olfactory loss for human romantic and sexual activity. *Archives of Sexual Behavior* 49:1295–1302. doi: 10.1007/s10508-020-01664-2
26. Pernet, C. R., McAleer, P., Latinus, M., et al. (2015). The human voice areas: Anatomical and functional organization. *Cerebral Cortex*, 25, 4656–4669.
27. Richoz AR, et al. (2017). Audio-visual perception of gender by infants emerges earlier for female than male faces. *Frontiers in Psychology* 8:149.
28. Salu Y. (2011). The roots of sexual arousal and sexual orientation. *Medical Hypotheses* 76:384–387. doi: 10.1016/j.mehy.2010.10.048
29. Salu Y. (2020). Nature controls nurture in the development of sexual orientation, and voice is nature’s agent. Preprints. doi: 10.20944/preprints202003.0134.v1
30. Savic I, Berglund H, Gulyas B and Roland P. (2001). Smelling of odorous sex hormone-like compounds causes sex-differentiated hypothalamic activations. *Neuron* 31:661–668.
31. Savic I and Lindström P. (2008). PET and MRI show differences in cerebral asymmetry and functional connectivity between homosexual and heterosexual subjects. *Proceedings of the National Academy of Sciences of the USA* 105:9403–9408.
32. Saxton, T. K., Lyndon, A., Little, A. C., & Roberts, S. C. (2008). Evidence for sex pheromones in humans? *Proceedings of the Royal Society B*, 275, 2723–2729.
33. Sisk CL and Zehr JL. (2005). Pubertal hormones organize the adolescent brain and behavior. *Frontiers in Neuroendocrinology* 26:163–174.
34. Sokhi DS, Hunter MD, Wilkinson ID and Woodruff PWR. (2005). Male and female voices activate distinct regions in the male brain. *NeuroImage* 27:572–578. doi: 10.1016/j.neuroimage.2005.04.023
35. Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR and Dale AM. (2004). Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *Journal of Neurophysiology* 91:1282–1296. doi: 10.1152/jn.01125.2002

36. Trotier D. (2011). Vomeronasal organ and human pheromones. *European Annals of Otorhinolaryngology, Head and Neck Diseases* 128:184–190. doi: 10.1016/j.anorl.2010.11.008
37. Yu ME, Johnson TJ and Fecher N. (2024). Learning to identify talkers: Do 4.5-month-old infants recognize unfamiliar female talkers? *JASA Express Letters* 4:015203.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.