

UNSOLVED MYSTERY

How did vocal communication come to dominate human language? A view from the womb

Alexis Hervais-Adelman ^{1,2,3*}, Simon W. Townsend^{3,4}

1 Department of Basic Neuroscience, University of Geneva, Geneva, Switzerland, **2** Zurich Centre for Linguistics, University of Zurich, Zurich, Switzerland, **3** Institute for the Interdisciplinary Study of Language Evolution, University of Zurich, Zurich, Switzerland, **4** Department of Evolutionary Anthropology, University of Zurich, Zurich, Switzerland

* alexis.hervais-adelman@unige.ch



Abstract

Whether human language evolved via a gestural or a vocal route remains an unresolved and contentious issue. Given the existence of two preconditions—a “language faculty” and the capacity for imitative learning both vocally and manually—there is no compelling evidence for gesture being inherently inferior to vocalization as a mode of linguistic expression; indeed, signed languages are capable of the same expressive range as spoken ones. Here, we revisit this conundrum, championing recent methodological advances in human neuroimaging (specifically, in utero functional magnetic resonance imaging) as a window into the role of the prenatal gestational period in language evolution, a critical, yet currently underexplored environment in which fetuses are exposed to, and become attuned to, spoken language. In this Unsolved Mystery, we outline how, compared to visual sensitivity, the ontogenically earlier development of auditory sensitivity, beginning in utero and persisting for several months *post-partum*, alongside the relative permeability of the uterine environment to sound, but not light, may constitute a small but significant contribution to the current dominance of spoken language.

OPEN ACCESS

Citation: Hervais-Adelman A, Townsend SW (2025) How did vocal communication come to dominate human language? A view from the womb. *PLoS Biol* 23(4): e3003141. <https://doi.org/10.1371/journal.pbio.3003141>

Published: April 15, 2025

Copyright: © 2025 Hervais-Adelman. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: AH-A and SWT were supported by the Swiss National Science Foundation, <http://www.snf.ch>; grant: PP00P1_206155, 100014_204299 and 225146 to AH-A and 198912 and 225146 to SWT. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: GA, gestational age; NICU, neonatal intensive care unit.

Introduction

Human expressive language is intrinsically biased toward the spoken modality, yet how speech came to be the primary mode of communication remains unclear. One dominant explanation is that language evolved in the gestural modality with spoken language then piggybacking on this existing cognitive architecture and ultimately becoming the dominant communicative modality [1–3]. Classical arguments in favor of gestural evolution of language suggest that, since great apes exhibit more flexible communication in the visual-gestural than vocal modality, our ancestors would have also been gestural communicators [4,5]. An additional argument suggests that in humans, gestural communication begins earlier in infants [6] and, under the premise

that ontogeny recapitulates phylogeny, a similar communicative progression may have occurred over evolutionary time.

Whilst perhaps intuitive, the gestural origins hypothesis has recently received substantial challenges from both developmental and comparative perspectives. Firstly, a systematic quantification of the rates of speech-like vocalizations and gestures over early infancy suggests that at 3 months infants produce 35 times more vocal protophonetic than gestural acts [7], indicating that vocalizations, not gestures, are the primary mode of communication in infancy. Secondly, while great apes do indeed deploy their gestures flexibly, this is also true of their vocalizations, and a growing body of evidence suggests that primate vocal communication arguably displays more similarities with key components of human language than gestures [8,9].

These shortcomings of the gesture-first hypothesis are further compounded by the unresolved issue of the precise catalyst promoting a shift from gestural to vocal communication. Numerous benefits of speech over gesture have been proposed, ranging from facilitating communication in the dark or at a distance, to predator avoidance and the potential for communication while leaving the hands free for tool use or infant carrying. Although it cannot be disputed that vocal communication does have these properties, and that they are likely beneficial, an argument that they can drive the evolution of the dominance of the vocal mode has received little compelling empirical support. The practical benefits to the communicative modality cannot serve as explanations for the selection of speech over gesture. Evolution, as it is currently understood, does not proceed on the basis of “possible but as-yet-unrealized advantages” [10]. Rather, changes arise cumulatively, on the basis of gradual shifts that confer benefits for contemporaneous functions. It is difficult, if not impossible, to establish from the existing (fossil) record or from phylogenetic comparisons that the presumed incremental changes from gesture to vocalization was indeed sufficiently adaptive at intermediate stages to drive a selection benefit in its favor.

Here, we argue that the debate so far has neglected to integrate a class of clues from ontogeny that can cast light on the driving factors that have been instrumental in promoting the dominance of the vocal channel. Specifically, we contend that the nature and constraint of the prenatal period, i.e., gestation, play a powerful role in this story. We compare the development and maturation of the auditory and visual systems in utero and outline the evidence in favor of a potent role for acoustic information in tuning a newborn’s *post-partum* behavior. Evidence from emerging signed languages in deaf communities [11] and communities with a high proportion of congenitally deaf members [12] indicates that the necessary preconditions are in place for the development of vocal or gestural expressive language given the appropriate conditions for cultural transmission after birth [13,14]. In addressing the unsolved mystery of why the vocal mode has come to dominate human communication despite the availability of two viable channels, we contend that the combination of the environment in utero and the development of fetal sensory and cognitive capacities during the third trimester of gestation is such that it can support environmentally driven audiomotor learning to a significantly greater degree than visuo-gestural.

The different ontogenetic timelines of hearing and vision

Our basic premise is that the uterine environment is permissive to auditory signals, and that the human auditory system begins to be responsive to sound at approximately 6 months of gestational age (GA) [15] and increases in its range of frequency sensitivity during gestation [16]. This provides the necessary preconditions for a human fetus to be materially affected by auditory stimuli. In contrast, the womb is poorly penetrant to light, and the human visual system suffers from very poor acuity for the first months of life. The head-start therefore available to the development of auditory-vocal over visuo-gestural communication may constitute an additional significant driving factor in tilting the balance in favor of a vocal mode. Moreover, this relative developmental lag for vision is not confined to the uterine environment—even when the maternal abdomen no longer blocks visual input *post-partum*, the scope for learning about gestures is limited by the acuity of the visual system, which does not reach adult-like levels until between 24 and 36 months *post-partum* (see Fig 1) [17]. All the while, the peripheral auditory system is sensitive and responsive. While it is not feasible to reliably determine neonatal auditory sensitivity from overt behavioral measures, data using auditory brainstem responses suggest that the auditory system is responsive across the frequency range 0.5–8 kHz, albeit with moderately elevated absolute thresholds compared to adults [18]. The vocal channel therefore benefits from a 3-month in utero head-start over the gestural, in addition to a substantial *post-partum* interval during which the visual system provides only poor resolution. We argue that this provides ample basis for a precocious bias of the habitual communication channel in favor of the auditory-vocal over the visuo-gestural.

The fetal auditory system responds to sound, and extracts and stores regularities

The fetal environment is very different to the adult environment in which acoustic vibrations are transmitted to the ear through air. The fetus develops in a fluid-filled sac inside the mother's abdomen that shields the fetus from the outside

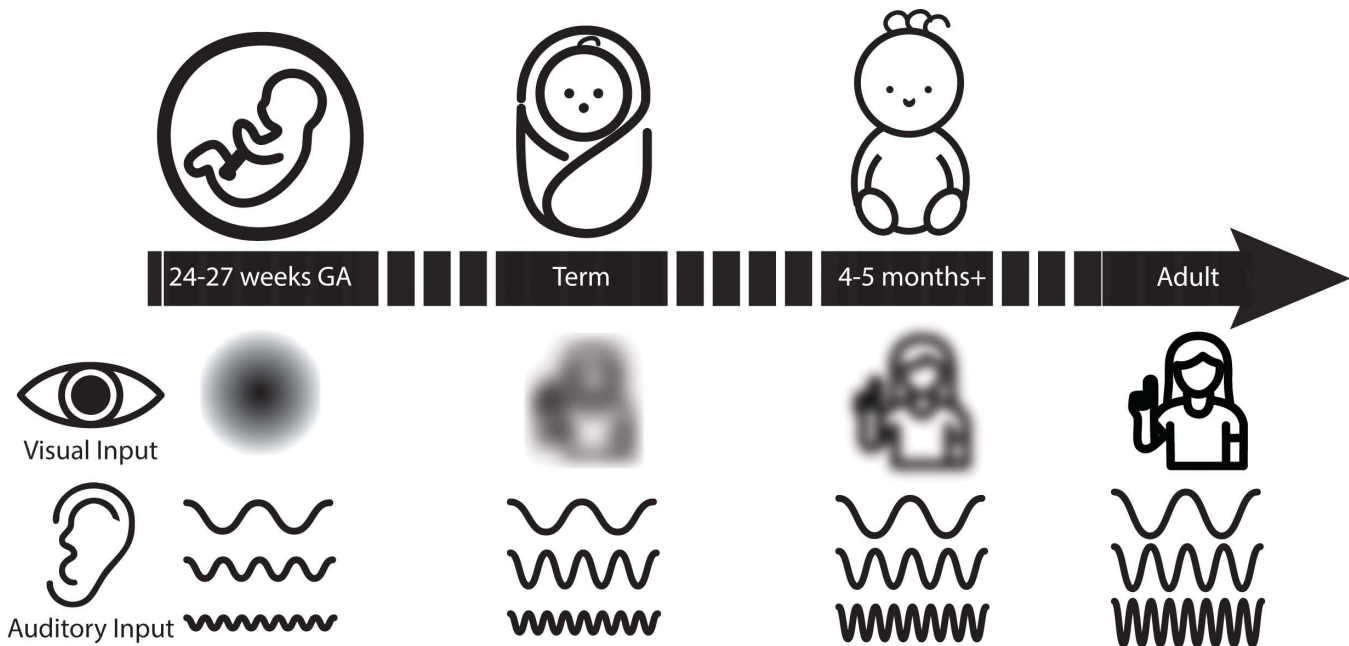


Fig 1. A schematic representation of the timeline of fetal visual and auditory perception. In utero, acoustic stimulation is available, while only luminosity is available to the visual system. At birth, although the physical barrier of the abdomen is no longer in place, the acuity of the fetal visual system is very poor. At several months *post-partum*, visual acuity has improved but remains poorer than that of adults until approximately 6 months. Meanwhile, the auditory system, despite having relatively elevated thresholds compared to adults, especially in higher frequencies, is nevertheless responsive to a broad range of frequencies (see text for details) from approximately 24 gestational weeks onward.

<https://doi.org/10.1371/journal.pbio.3003141.g001>

world and also has an attenuating effect on the acoustical information that can reach the fetal ear. Nevertheless, a considerable body of work suggests that all frequencies within the typical human auditory range can be transmitted to the fetus, albeit with an attenuation on the order of 25–30 dB [19]. This implies that all maternal speech and any sufficiently loud speech of others will be transmitted to the fetal cochlea. One study [20] claims that although only approximately 30% of phonetic information from speech in the environment is available to the fetal auditory system, intonation is very well transmitted. The uterine environment features a constant background noise of maternal sounds, particularly the heartbeat and digestive processes, but these are confined to lower frequencies (below 300 Hz) [21] and thus do not mask all other sounds. Recent work using hydrophones affixed to the occiput of fetal sheep in utero has revealed that the transmission of sound from 100 Hz to 20 kHz is moderately attenuated by the uterine environment in the amniotic sac [22], but to a lesser degree with less high-frequency roll-off than previously thought.

Empirical evidence indicates that the human fetus has a fully functioning auditory system at 6 months GA [15], suggesting that the fetus should be able to perceive and learn from auditory stimuli in the womb. Indeed, there is overwhelming evidence that fetal auditory learning occurs both in non-linguistic and linguistic auditory domains and is retained *post-partum*. Newborn infants discriminate between their mother's voice and that of other unknown female speakers [23], between their native language and other languages [24,25], and they even exhibit preferences for melodies [26] or stories [27] to which they were exposed in utero. Further, it has been shown that newborn infants discriminate native and foreign vowels [28], and that they do so in a manner consistent with having begun to develop native vowel categories in utero. Other evidence points to significant in utero learning of the properties of trisyllabic nonwords [29] reflected in *post-partum* neural traces consistent with advanced sequence recognition and pitch discrimination capabilities. It should be noted that testing newborns or infants a few days *post-partum* may carry the risk of some influence of post-natal learning. This is not, however, the case for evidence for auditory sensitivity and habituation that has been acquired in utero. For example, fetal heart rate changes have been recorded in response to airborne auditory stimuli demonstrating patterns consistent with habituation and dishabituation [30]. Even earlier work [31] had already demonstrated fetal heart-rate changes consistent with a discriminative ability at 35 weeks GA for airborne syllable presentation. Further, evidence from in utero magnetoencephalography, a non-invasive technique that measures magnetic field fluctuations induced by neuronal activity, also suggests that the fetal brain produces a classical mismatch response when presented with a novel stimulus in a train of identical stimuli [32,33]. Thus, prior to birth, the fetus processes and derives regularities from the auditory input, laying the foundations for future recognition, categorization and, ultimately, comprehension of speech (see Fig 2). We therefore contend that early experience with speech stimuli, both environmental and maternal in origin has the potential to exert developmental influences on the in utero brain.

Methodological advances in MRI enable finer-grained insights into the fetal brain

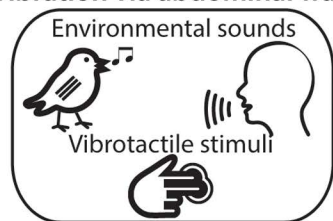
Recent advances in neuroimaging techniques and growing acceptance of the technique have facilitated the take-up of fetal neuroimaging as an investigative (not merely clinical) technique. In utero neuroimaging is a valuable source of information for revealing the structural and functional organization of the fetal brain. Over the past decades, in utero functional imaging has demonstrated that fetal brains respond to acoustic stimulation in auditory cortical areas [34], and display differentiated cerebral responses to the maternal *versus* non-maternal voice [35–37]. Structural and resting-state imaging of fetal brains also provides new insights into the underlying organization of the brain as a whole, and specifically in speech and language-related networks.

The available data on in utero development of auditory and speech circuits has recently been comprehensively reviewed [38] and there is overwhelming evidence that not only does the auditory system respond to stimuli from around six months GA, it is embedded in an increasingly complex fronto-temporal network developing from 26–27 weeks GA onwards [39,40], and that these networks become increasingly adult-like. For example, one study [41] showed a 61.66% overlap between adult and fetal connectomes in the third trimester.

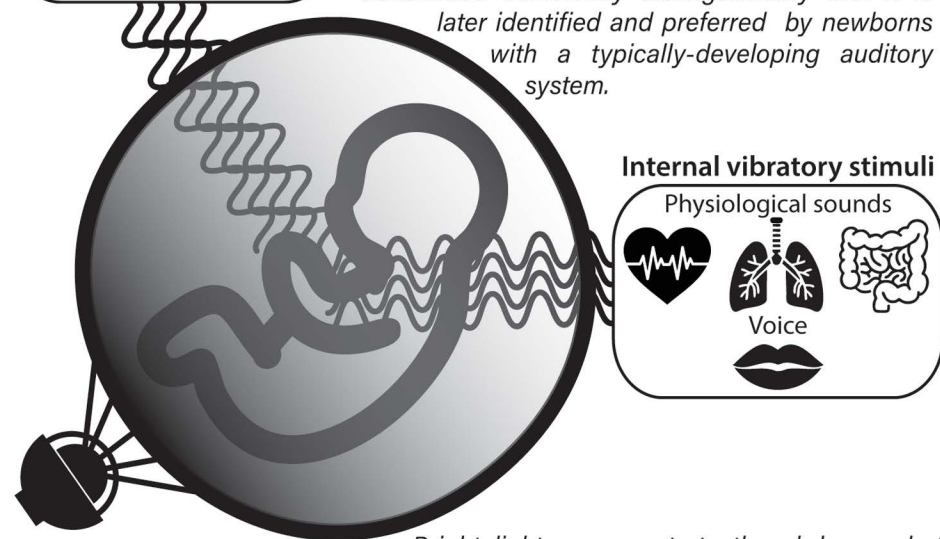
Prenatal sounds can alter developmental trajectories

Significantly, evidence abounds that prenatal auditory sensitivity provides a potential mechanism for “auditory developmental programming”. It has now been demonstrated that prenatal sounds provoke phenotypic changes in embryos across diverse taxa, including invertebrates, amphibians, reptiles, and birds [comprehensively reviewed in 42], essentially demonstrating that there is a mechanism for tailoring the fitness of an individual to the environment in response to an external acoustic (or vibratory) signal. A number of striking examples of this phenomenon exist. For example, in invertebrates, the synchronized hatching of stinkbugs (*Halyomorpha halys*) in response to an egg-cracking vibration causes all members of a batch to emerge simultaneously [43], with a proposed function of avoiding egg cannibalism by hatchlings. Shield bugs (*Parastrachia japonensis*) are attested to hatch synchronously in response to maternal vibrations [44]. Among amphibians, red-eyed tree frogs (*Agalychnis callidryas*) avoid egg predation by responding to vibrations from approaching predators with immediate early hatching [45], as do multiple species of glass frogs [46]. In birds, multiple examples of *in*

Vibration via abdominal wall



Sound from both exogeneous and endogeneous sources penetrates the womb and is detected by the fetal auditory system, driving auditory learning that persists post-partum. Exogeneous speech stimuli are perceived with sufficient fidelity for in utero vowel discrimination. The maternal voice is transmitted sufficiently distinguishably that it is later identified and preferred by newborns with a typically-developing auditory system.



Light via abdominal wall

Bright light can penetrate the abdomen, but transmission is impaired by tissue. The fetus is exposed to changes in luminance but not to any clear images. Combined with poor visual acuity, opportunities for extracting patterns from the visual world are thus significantly limited.

Fig 2. Schematic illustration of sensory input to the fetus. The diagram highlights the relative penetrance of vibrotactile and acoustic stimuli, which can be transmitted directly as pressure waves, compared to light, whose photons are absorbed by the abdomen, reflected, and scattered, preventing the transmission of images. All icons are original works or reused under a CC-BY4.0 license. “Vibrotactile stimuli” icon obtained from <https://icon-icons.com/icon/touching-ngers/64385>; “heartbeat” icon by M2n obtained from <https://thenounproject.com/icon/heartbeat-5770628/>.

<https://doi.org/10.1371/journal.pbio.3003141.g002>

ovo learning have been reported, including the acquisition of a vocal “password” by superb fairy wren (*Malurus cyaneus*) chicks [47] from maternal incubation calls. Elements of the calls are reprised by the hatchlings in their begging calls, thereby identifying them as biological brood members, ensuring that they receive food, and reducing the risk of parasitism from cuckoo chicks unfamiliar with the song. There are even more potent experimental demonstrations of the extent to which prenatal sound exposure can alter individual phenotypes directly. In one recent demonstration, prenatal (*in ovo*) exposure to “heat-calls” in arid-adapted zebra finches was shown to drive development of superior heat-adaptation in mitochondrial metabolism [48]. In chickens, exposure to conspecific calls, and even music, *in ovo* stimulates neural and cognitive development resulting in enhanced spatial learning abilities in the days after hatching [49]. In utero learning of the maternal voice in humans is a manifestation of such a widespread biological phenomenon—it is presumably beneficial to the newborn infant to immediately orient to its mother as the likely primary caregiver immediately after birth.

Although the experimental evidence for such obvious functional adaptations in humans is relatively scarce, one compelling result is that there is a significant difference in the structural and functional development of the auditory system of pre-term infants compared to term infants. It is believed that this arises because of differential exposure to acoustic input, with pre-term infants being subjected to a substantially different auditory environment in a neonatal intensive care unit (NICU) compared to their term peers. It has been shown that exposure to recordings of the maternal voice and heartbeat during NICU stays leads to increased volume of primary auditory cortex at term compared to exposure only to standard hospital noise [50].

Prenatal programming of human communicative behavior

A puzzling finding in the literature on human prenatal learning is that newborn infants produce cries that carry the pitch accent (or “melody”) of the language of their gestational environment. This accent is presumably learned primarily via exposure to the maternal voice. This intriguing observation was first published by Mampe and colleagues [51] who conducted a study comparing the cries of newborn full-term infants who gestated in either a monolingual French or monolingual German environment. They demonstrated that the spontaneous cries that they recorded from these infants 2–5 days *post-partum* had the melodic contour typical of speech in the environmental language, leading to the remarkable conclusion that the environmental language experienced in utero shapes the acoustics of infants’ cries. These findings have since been replicated in other linguistic environments. Wermke and colleagues [52] demonstrated that the fundamental frequency of cries was more variable in newborns whose ambient language was Lamnso, a complex tonal language of the Niger-Congo family, than in German newborns. In another study, Wermke and colleagues [53] compared the melodic contours of neonates who gestated in monolingual German and Mandarin Chinese environments, and found that the cries of Chinese newborns were more melodically complex than those of their German counterparts. This is consistent with the melodic contour of Mandarin (a tonal language) being more complex than that of German. Similarly, Prochnow and colleagues [54] analyzed the cries of Swedish newborns, and contrasted them with German newborns. Swedish is a pitch-accent language, and German is not. Swedish newborns produced cries of greater melodic complexity. Gustafson and colleagues have queried the methodology of some of this work [55], arguing that whilst individual infants’ cries are distinctive, the evidence for language-group differences is unconvincing. Nevertheless, a recent study using a neural network approach to classification showed that it was possible to accurately classify the native language of infants born in Germany or Japan to monolingual speakers of German and Japanese, respectively, based on samples of cries recorded at 2–6 days after birth [56]. Although this may leave some margin for *post-partum* learning, if these effects generalize, they constitute an extremely precocious expression of language-induced modification of cries.

Taken together, these studies indicate that a linguistic bias is present in vocal behavior extremely early in the infant’s *post-partum* life. Specifically, there is an impact of the melodic contour of the environmental speech on infants’ earliest communicative gestures (for the proposed underlying mechanisms, see Box 1). This is manifested in, but overlaid upon, the universal innate reflex of crying.

Box 1. Substrates of audiomotor learning in the womb

The neuroanatomical basis for an auditory-motor mapping is provided by the arcuate fasciculus, a white matter bundle that connects posterior temporal areas to inferior frontal areas, via the parietal lobule [57], which is thought to be instrumental in vocal learning, by providing the means for auditory feedback to tune articulation [58]. Newborns have a precursor of the arcuate fasciculus, similarly connecting auditory areas to premotor and somatosensory areas [59–61]. Although there is scant characterization of this connection prior to birth, from in utero, preterm or *ex vivo* samples, it must be present prior to term, laying the cerebral ground for a direct influence of auditory stimulation on the motor system, potentially for the entire last trimester of pregnancy, once the auditory system is functional. However, in the uterine environment the fetus cannot receive auditory feedback about vocalizations; generating an audible vocalization requires the creation of a pressure differential between air in the lungs and air outside the lungs, forcing air through the vocal folds. In utero, the fetus is immersed in a fluid-filled environment and consequently cannot engage in audible vocal behavior, and thus receives no auditory feedback on articulatory gestures that may be executed. Thus, auditory feedback is unlikely to contribute to in utero tuning of articulatory motor patterns. However, one source of feedback that is available is somatosensory feedback. If the fetus were to execute articulatory motor gestures, in utero somatosensory feedback would likely be received. Since pitch-contour and amplitude of infant cries are highly correlated [62], it is possible that if the fetus were to learn only to match its motor output with the typical intensity contour of the environmental speech input, this would provide a basis for developing articulatory motor control patterns that bias cry production towards the native accent. It is possible that the extraction of patterns from auditory stimulation depends upon an amodal representation of magnitude over time that is stored and can be retrieved in the service of multiple domains. Even if this is the case, the fetus has limited scope for movement that may recapitulate these patterns, due to the relative paucity of space in the womb. Indeed, other than the digits, only orofacial movements can readily be executed, again potentially biasing the system towards vocal motor learning.

The potential significance of such accented crying remains a matter of speculation. However, the discovery of accents in newborns' cries has frequently been considered from the perspective of speech development, and indeed the melodic quality of infants' cries has been shown to be predictive of language outcomes in the second year of life [63]. This tantalizingly suggests that whatever mechanisms enable precocious acquisition of a "native-like" vocal-motor repertoire are of substantial importance to language development, and, significantly, suggest that crucial determinants of speech acquisition in play during gestation, manifest in cry melody. However, another important question relating to this has not been addressed—does crying with a melodic contour that reflects the typical pitch-contour of the environmental (and likely native) language confer a benefit to the infant immediately, rather than to the future toddler? Does this precocious bias in vocal production serve as a signal to potential caregivers, identifying the infant as a member of their (linguistically-defined) community? Could the superposition of a native-like melody over a cry render that cry a superior signal for eliciting and optimizing caregiver responses, by maximizing salience and reducing aversiveness?

In addition to this theoretical possibility, recent ultrasound investigations during the third trimester have revealed that the fetus produces lip movements in response to maternal speech [64,65], and that when these lip movements occur, they are consistent with syllables enunciated by the mother. This is of particular interest as it suggests that the fetus may engage in an articulatory mirroring of auditory input. The likelihood that this is automatic, rather than voluntary, is suggested by the fact that it is generally thought the fetus is not conscious [66]—any suggestion of deliberate practice should be regarded cautiously. This provides a basis for elaborating a theory of how in utero experience of speech can drive the adaptation of the articulatory motor apparatus to have a bias towards producing natively-accented utterances. Namely, that the fetus possesses a mechanism that automatically maps speech sounds from the auditory to the motor domain,

which contributes to developing articulatory-motor patterns executed at birth. Interestingly, similar data have been forwarded in non-human primates: a recent study [67] used ultrasound to demonstrate in utero practicing of vocal gestures that are to be produced after birth in marmosets. It should be noted that in both humans and non-human primates, whilst the work demonstrating these effects is limited in scale, and that further investigations are required to establish with certainty whether articulatory mimicry in utero is a reliable phenomenon, the presence of a vocal gestural repertoire (albeit a limited one) in utero provides an intriguing potential substrate for the influence of auditory input.

Solving the mystery?

The arguments presented above do not solve the mystery of the dominance of the vocal mode. However, there is an overwhelming mass of evidence suggesting that prenatal auditory experience alters the developing brain and that it can do so in a manner that is consistent with programming for a phenotypic benefit for an individual. The existence of such a well-conserved mechanism across species (and across taxonomic groups, including amphibians, reptiles, birds, and mammals) suggests that the influence of the uterine environment must be considered in the story of human language evolution. The nature of expressive language is constrained by the available channels, and it is likely that the prenatal influences that drive toward the development of a dominant vocal mode also result in a reciprocal influence of the mode on the underlying form. If a co-evolution of production and perception is assumed, it is also probable that these co-evolved with the relevant cognitive architecture for encoding and decoding linguistic messages, and that the output modality will have exercised some constraints shaping the nature of the system itself. We hope that approaching this question using the novel imaging tools at our disposal and with a focus on enquiring into the potential mechanisms and results of prenatal learning will provide a framework for novel modes of enquiry into the key drivers implicated in the evolution of language.

Although the refrain “further work is needed” is worn past the point of cliché, it remains as true as ever in addressing the question “why do humans speak?”. We hope future work will shed further light on the proximal benefits incurred by crying. Whilst challenging, large-scale behavioral experiments investigating preferences (e.g., inviting ratings of aversiveness and salience) of cries showing the melody of listeners’ native languages can provide a first insight into whether adults are sensitive to this dimension. Further studies, for example investigating autonomic physiological responses, such as galvanic skin response or pupillometry, could also provide a means for covert evaluation of whether cry melody (or other features of the acoustics of cries) can have a substantive impact on adult listeners. Neuroimaging studies of adult responses to cries having different melodies could also shed some light on the neurophysiological consequences of exposure to cries conveying this potential signal of group membership. Additional comparative work, particularly in our closest living primate relatives, will also be instrumental in helping further unpack the evolutionary emergence of prenatal learning and the extent to which such abilities are unique to humans or rooted more deeply in the primate lineage. Finally, thanks to advances in non-invasive imaging technologies, in particular fetal fMRI, it is now possible to explore the cerebral bases of in utero perception and prenatal learning, which may bring us one step closer to understanding the impact of exposure to a rich auditory environment on the brain and its consequences for communicative behavior. Ultimately, it seems that elucidating the foundations of speech will require a concerted effort to understand what is happening even before birth.

Author contributions

Conceptualization: Alexis Hervais-Adelman.

Funding acquisition: Alexis Hervais-Adelman, Simon W. Townsend.

Project administration: Alexis Hervais-Adelman.

Writing – original draft: Alexis Hervais-Adelman, Simon W. Townsend.

Writing – review & editing: Alexis Hervais-Adelman, Simon W. Townsend.

References

1. Corballis MC. From hand to mouth: the origins of language. Princeton University Press; 2002
2. Corballis MC. Language evolution: a changing perspective. *Trends Cogn Sci*. 2017;21(4):229–36. <https://doi.org/10.1016/j.tics.2017.01.013> PMID: [28214132](https://pubmed.ncbi.nlm.nih.gov/28214132/)
3. Sterelny K. Language, gesture, skill: the co-evolutionary foundations of language. *Philos Trans R Soc Lond B Biol Sci*. 2012;367(1599):2141–51. <https://doi.org/10.1098/rstb.2012.0116> PMID: [22734057](https://pubmed.ncbi.nlm.nih.gov/22734057/)
4. Call J, Tomasello M. The gestural communication of apes and monkeys. Psychology Press; 2020.
5. Tomasello M, Call J. Thirty years of great ape gestures. *Anim Cogn*. 2019;22(4):461–9.
6. Caselli MC, Rinaldi P, Stefanini S, Volterra V. Early action and gesture “vocabulary” and its relation with word comprehension and production. *Child Dev*. 2012;83(2):526–42. <https://doi.org/10.1111/j.1467-8624.2011.01727.x> PMID: [22304431](https://pubmed.ncbi.nlm.nih.gov/22304431/)
7. Burkhardt-Reed MM, Long HL, Bowman DD, Bene ER, Oller DK. The origin of language and relative roles of voice and gesture in early communication development. *Infant Behav Dev*. 2021;65:101648. <https://doi.org/10.1016/j.infbeh.2021.101648> PMID: [34628105](https://pubmed.ncbi.nlm.nih.gov/34628105/)
8. Leroux M, Schel AM, Wilke C, Chandia B, Zuberbühler K, Slocombe KE, et al. Call combinations and compositional processing in wild chimpanzees. *Nat Commun*. 2023;14(1):2225. <https://doi.org/10.1038/s41467-023-37816-y> PMID: [37142584](https://pubmed.ncbi.nlm.nih.gov/37142584/)
9. Leroux M, Townsend SW. Call combinations in great apes and the evolution of syntax. *Anim Behav Cogn*. 2020;7(2):131–9.
10. Kendon A. Reflections on the “gesture-first” hypothesis of language origins. *Psychon Bull Rev*. 2017;24(1):163–70. <https://doi.org/10.3758/s13423-016-1117-3> PMID: [27439503](https://pubmed.ncbi.nlm.nih.gov/27439503/)
11. Senghas A, Coppola M. Children creating language: how Nicaraguan sign language acquired a spatial grammar. *Psychol Sci*. 2001;12(4):323–8. <https://doi.org/10.1111/1467-9280.00359> PMID: [11476100](https://pubmed.ncbi.nlm.nih.gov/11476100/)
12. Ergin R. Central Taurus sign language: a unique vantage point into language emergence [Ph.D.]. United States—Massachusetts: Tufts University; 2017.
13. Lenneberg EH. The biological foundations of language. *Hosp Pract*. 1967;2(12):59–67.
14. Sandler W. Speech and sign: the whole human language. *Theor Linguist*. 2024;50(1–2):107–24. <https://doi.org/10.1515/tl-2024-2008>
15. Birnholz JC, Benacerraf BR. The development of human fetal hearing. *Science*. 1983;222(4623):516–8. <https://doi.org/10.1126/science.6623091> PMID: [6623091](https://pubmed.ncbi.nlm.nih.gov/6623091/)
16. Hepper PG, Shahidullah BS. Development of fetal hearing. *Arch Dis Child Fetal Neonatal Ed*. 1994;71(2):F81–7. <https://doi.org/10.1136/fn.71.2.f81> PMID: [7979483](https://pubmed.ncbi.nlm.nih.gov/7979483/)
17. Mayer DL, Dobson V. Visual acuity development in infants and young children as assessed by operant preferential looking. *Vision Res*. 1982;22(9):1141–51. [https://doi.org/10.1016/0042-6989\(82\)90079-7](https://doi.org/10.1016/0042-6989(82)90079-7) PMID: [7147725](https://pubmed.ncbi.nlm.nih.gov/7147725/)
18. Sininger YS, Abdala C. Hearing threshold as measured by auditory brain stem response in human neonates. *Ear Hearing*. 1996;17(5):395–401.
19. Granier-Deferre C, Ribeiro A, Jacquet A-Y, Bassereau S. Near-term fetuses process temporal features of speech. *Dev Sci*. 2011;14(2):336–52. <https://doi.org/10.1111/j.1467-7687.2010.00978.x> PMID: [22213904](https://pubmed.ncbi.nlm.nih.gov/22213904/)
20. Querleu D, Renard X, Versyp F, Paris-Delrue L, Crèpin G. Fetal hearing. *Eur J Obstet Gynecol Reprod Biol*. 1988;28(3):191–212. [https://doi.org/10.1016/0028-2243\(88\)90030-5](https://doi.org/10.1016/0028-2243(88)90030-5) PMID: [3061844](https://pubmed.ncbi.nlm.nih.gov/3061844/)
21. Abrams RM, Gerhardt KJ. The acoustic environment and physiological responses of the fetus. *J Perinatol*. 2000;20(8 Pt 2):S31–6. <https://doi.org/10.1038/sj.jp.7200445> PMID: [11190698](https://pubmed.ncbi.nlm.nih.gov/11190698/)
22. Gelat P, David AL, Haqhenas SR, Henriques J, Thibaut de Maisieres A, White T, et al. Evaluation of fetal exposure to external loud noise using a sheep model: quantification of in utero acoustic transmission across the human audio range. *Am J Obstet Gynecol*. 2019;221(4):343.e1–e11.
23. DeCasper AJ, Fifer WP. Of human bonding: newborns prefer their mothers’ voices. *Science*. 1980;208(4448):1174–6. <https://doi.org/10.1126/science.7375928> PMID: [7375928](https://pubmed.ncbi.nlm.nih.gov/7375928/)
24. Mehler J, Jusczyk P, Lambertz G, Halsted N, Bertocini J, Amiel-Tison C. A precursor of language acquisition in young infants. *Cognition*. 1988;29(2):143–78. [https://doi.org/10.1016/0010-0277\(88\)90035-2](https://doi.org/10.1016/0010-0277(88)90035-2) PMID: [3168420](https://pubmed.ncbi.nlm.nih.gov/3168420/)
25. Moon C, Cooper RP, Fifer WP. Two-day-olds prefer their native language. *Infant Behav Dev*. 1993;16(4):495–500. [https://doi.org/10.1016/0163-6383\(93\)80007-u](https://doi.org/10.1016/0163-6383(93)80007-u)
26. Hepper PG. An examination of fetal learning before and after birth. *Irish J Psychol*. 1991;12(2):95–107.
27. DeCasper AJ, Spence MJ. Prenatal maternal speech influences newborns perception of speech sounds. *Infant Behav Dev*. 1986;9(2):133–50.
28. Moon C, Lagercrantz H, Kuhl PK. Language experienced in utero affects vowel perception after birth: a two-country study. *Acta Paediatr*. 2013;102(2):156–60. <https://doi.org/10.1111/apa.12098> PMID: [23173548](https://pubmed.ncbi.nlm.nih.gov/23173548/)
29. Partanen E, Kujala T, Naatanen R, Liitola A, Sambeth A, Huotilainen M. Learning-induced neural plasticity of speech processing before birth. *Proc Natl Acad Sci U S A*. 2013;110(37):15145–50. <https://doi.org/10.1073/pnas.1302159110> PMID: [23980148](https://pubmed.ncbi.nlm.nih.gov/23980148/)
30. Kisilvesky BS, Muir DW. Human fetal and subsequent newborn responses to sound and vibration. *Infant Behav Dev*. 1991;14(1):1–26. [https://doi.org/10.1016/0163-6383\(91\)90051-s](https://doi.org/10.1016/0163-6383(91)90051-s)

31. Lecanuet JP, Granier-Deferre C, DeCasper AJ, Maugeais R, Andrieu AJ, Busnel MC. Fetal perception and discrimination of speech stimuli as demonstrated by cardiac reactivity preliminary results. *C R Acad Sci III*. 1987;305(5):161–4. PMID: [3113681](https://pubmed.ncbi.nlm.nih.gov/3113681/)
32. Muenssinger J, Matuz T, Schleger F, Draganova R, Weiss M, Kiefer-Schmidt I, et al. Sensitivity to auditory spectral width in the fetus and infant—an fMEG study. *Front Hum Neurosci*. 2013;7:917. <https://doi.org/10.3389/fnhum.2013.00917> PMID: [24427130](https://pubmed.ncbi.nlm.nih.gov/24427130/)
33. Muenssinger J, Matuz T, Schleger F, Kiefer-Schmidt I, Goelz R, Wacker-Gussmann A, et al. Auditory habituation in the fetus and neonate: an fMEG study. *Dev Sci*. 2013;16(2):287–95. <https://doi.org/10.1111/desc.12025> PMID: [23432837](https://pubmed.ncbi.nlm.nih.gov/23432837/)
34. Moore RJ, Vadeyar S, Fulford J, Tyler DJ, Gribben C, Baker PN, et al. Antenatal determination of fetal brain activity in response to an acoustic stimulus using functional magnetic resonance imaging. *Hum Brain Mapp*. 2001;12(2):94–9. [https://doi.org/10.1002/1097-0193\(200102\)12:2<94::aid-hbm1006>3.0.co;2-e](https://doi.org/10.1002/1097-0193(200102)12:2<94::aid-hbm1006>3.0.co;2-e) PMID: [11169873](https://pubmed.ncbi.nlm.nih.gov/11169873/)
35. Goldberg E, McKenzie CA, de Vrijer B, Eagleson R, de Ribaupierre S. Fetal response to a maternal internal auditory stimulus. *J Magn Reson Imaging*. 2020.
36. Jardri R, Houfflin-Debarge V, Delion P, Pruvo J-P, Thomas P, Pins D. Assessing fetal response to maternal speech using a noninvasive functional brain imaging technique. *Int J Dev Neurosci*. 2012;30(2):159–61. <https://doi.org/10.1016/j.ijdevneu.2011.11.002> PMID: [22123457](https://pubmed.ncbi.nlm.nih.gov/22123457/)
37. Jardri R, Pins D, Houfflin-Debarge V, Chaffiotte C, Rocourt N, Pruvo J-P, et al. Fetal cortical activation to sound at 33 weeks of gestation: a functional MRI study. *Neuroimage*. 2008;42(1):10–8. <https://doi.org/10.1016/j.neuroimage.2008.04.247> PMID: [18539048](https://pubmed.ncbi.nlm.nih.gov/18539048/)
38. Ghio M, Cara C, Tettamanti M. The prenatal brain readiness for speech processing: a review on foetal development of auditory and primordial language networks. *Neurosci Biobehav Rev*. 2021;128:709–19. <https://doi.org/10.1016/j.neubiorev.2021.07.009> PMID: [34274405](https://pubmed.ncbi.nlm.nih.gov/34274405/)
39. Jakab A, Schwartz E, Kasprian G, Gruber GM, Prayer D, Schöpf V, et al. Fetal functional imaging portrays heterogeneous development of emerging human brain networks. *Front Hum Neurosci*. 2014;8:852. <https://doi.org/10.3389/fnhum.2014.00852> PMID: [25374531](https://pubmed.ncbi.nlm.nih.gov/25374531/)
40. Thomason ME, Brown JA, Dassanayake MT, Shastri R, Marusak HA, Hernandez-Andrade E, et al. Intrinsic functional brain architecture derived from graph theoretical analysis in the human fetus. *PLoS One*. 2014;9(5):e94423. <https://doi.org/10.1371/journal.pone.0094423> PMID: [24788455](https://pubmed.ncbi.nlm.nih.gov/24788455/)
41. Turk E, van den Heuvel MI, Benders MJ, de Heus R, Franx A, Manning JH, et al. Functional connectome of the fetal brain. *J Neurosci*. 2019;39(49):9716–24. <https://doi.org/10.1523/JNEUROSCI.2891-18.2019> PMID: [31685648](https://pubmed.ncbi.nlm.nih.gov/31685648/)
42. Mariette MM, Clayton DF, Buchanan KL. Acoustic developmental programming: a mechanistic and evolutionary framework. *Trends Ecol Evol*. 2021;36(8):722–36. <https://doi.org/10.1016/j.tree.2021.04.007> PMID: [34052045](https://pubmed.ncbi.nlm.nih.gov/34052045/)
43. Endo J, Takanashi T, Mukai H, Numata H. Egg-cracking vibration as a cue for stink bug siblings to synchronize hatching. *Curr Biol*. 2019;29(1):143–8.e2.
44. Mukai H, Hironaka M, Tojo S, Nomakuchi S. Maternal vibration: an important cue for embryo hatching in a subsocial shield bug. *PLoS One*. 2014;9(1):e87932. <https://doi.org/10.1371/journal.pone.0087932> PMID: [24498224](https://pubmed.ncbi.nlm.nih.gov/24498224/)
45. Jung J, Kim SJ, Pérez Arias SM, McDaniel JG, Warkentin KM. How do red-eyed treefrog embryos sense motion in predator attacks? Assessing the role of vestibular mechanoreception. *J Exp Biol*. 2019;222(21):jeb.206052. <https://doi.org/10.1242/jeb.206052> PMID: [31586019](https://pubmed.ncbi.nlm.nih.gov/31586019/)
46. Delia J, Rivera-Ordóñez JM, Salazar-Nicholls MJ, Warkentin KM. Hatching plasticity and the adaptive benefits of extended embryonic development in glassfrogs. *Evol Ecol*. 2019;33:37–53. <https://doi.org/10.1007/s10682-018-9963-2>
47. Colombelli-Négrel D, Hauber ME, Robertson J, Sulloway FJ, Hoi H, Griggio M, et al. Embryonic learning of vocal passwords in superb fairy-wrens reveals intruder cuckoo nestlings. *Curr Biol*. 2012;22(22):2155–60. <https://doi.org/10.1016/j.cub.2012.09.025> PMID: [23142041](https://pubmed.ncbi.nlm.nih.gov/23142041/)
48. Udino E, George JM, McKenzie M, Pessato A, Crino OL, Buchanan KL, et al. Prenatal acoustic programming of mitochondrial function for high temperatures in an arid-adapted bird. *Proc R Soc B*. 2021;288(1964):20211893. <https://doi.org/10.1098/rspb.2021.1893> PMID: [34875198](https://pubmed.ncbi.nlm.nih.gov/34875198/)
49. Kauser H, Roy S, Pal A, Sreenivas V, Mathur R, Wadhwa S, et al. Prenatal complex rhythmic music sound stimulation facilitates postnatal spatial learning but transiently impairs memory in the domestic chick. *Dev Neurosci*. 2011;33(1):48–56. <https://doi.org/10.1159/000322449> PMID: [21212638](https://pubmed.ncbi.nlm.nih.gov/21212638/)
50. Webb AR, Heller HT, Benson CB, Lahav A. Mother's voice and heartbeat sounds elicit auditory plasticity in the human brain before full gestation. *Proc Natl Acad Sci U S A*. 2015;112(10):3152–7. <https://doi.org/10.1073/pnas.1414924112> PMID: [25713382](https://pubmed.ncbi.nlm.nih.gov/25713382/)
51. Mampe B, Friederici AD, Christophe A, Wermke K. Newborns' cry melody is shaped by their native language. *Curr Biol*. 2009;19(23):1994–7.
52. Wermke K, Teiser J, Yovsi E, Kohlenberg PJ, Wermke P, Robb M, et al. Fundamental frequency variation within neonatal crying: does ambient language matter? *Speech Lang Hearing*. 2016;19(4):211–7. <https://doi.org/10.1080/2050571x.2016.1187903>
53. Wermke K, Ruan Y, Feng Y, Dobnig D, Stephan S, Wermke P, et al. Fundamental frequency variation in crying of Mandarin and German neonates. *J Voice*. 2017;31(2):255.e25–e30. <https://doi.org/10.1016/j.jvoice.2016.06.009> PMID: [27397111](https://pubmed.ncbi.nlm.nih.gov/27397111/)
54. Prochnow A, Erlandsson S, Hesse V, Wermke K. Does a 'musical' mother tongue influence cry melodies? A comparative study of Swedish and German newborns. *Musicae Scientiae*. 2017;23(2):143–56.
55. Gustafson GE, Sanborn SM, Lin H-C, Green JA. Newborns' cries are unique to individuals (but not to language environment). *Infancy*. 2017;22(6):736–47.
56. Polzehl T, Herzig T, Wicke F, Wermke K, Khamsehashari R, Dahlem M, et al. Towards Classifying Mother Tongue from Infant Cries-Findings Substantiating Prenatal Learning Theory. *Proc Interspeech 2024*. 4199–203.
57. Catani M, Thiebaut de Schotten M. A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*. 2008;44(8):1105–32.

58. Hickok G, Poeppel D. The cortical organization of speech processing. *Nat Rev Neurosci.* 2007;8(5):393–402. <https://doi.org/10.1038/nrn2113> PMID: [17431404](https://pubmed.ncbi.nlm.nih.gov/17431404/)
59. Friederici AD The brain basis of language processing: from structure to function. *Physiol Rev.* 2011;91(4):1357–92.
60. Friederici AD. Evolution of the neural language network. *Psychon Bull Rev.* 2017;24(1):41–7. <https://doi.org/10.3758/s13423-016-1090-x> PMID: [27368631](https://pubmed.ncbi.nlm.nih.gov/27368631/)
61. Wilkinson M, Lim AR, Cohen AH, Galaburda AM, Takahashi E. Detection and growth pattern of arcuate fasciculus from newborn to adult. *Front Neurosci.* 2017;11:389.
62. Wermke K, Mende W, Manfredi C, Bruscoliglioni P. Developmental aspects of infant's cry melody and formants. *Med Eng Phys.* 2002;24(7–8):501–14. [https://doi.org/10.1016/s1350-4533\(02\)00061-9](https://doi.org/10.1016/s1350-4533(02)00061-9) PMID: [12237046](https://pubmed.ncbi.nlm.nih.gov/12237046/)
63. Wermke K, Leising D, Stellzig-Eisenhauer A. Relation of melody complexity in infants' cries to language outcome in the second year of life: a longitudinal study. *Clin Linguist Phon.* 2007;21(11–12):961–73. <https://doi.org/10.1080/02699200701659243> PMID: [17972192](https://pubmed.ncbi.nlm.nih.gov/17972192/)
64. Ferrari GA, Nicolini Y, Demuru E, Tosato C, Hussain M, Scesa E, et al. Ultrasonographic investigation of human fetus responses to maternal communicative and non-communicative stimuli. *Front Psychol.* 2016;7:354. <https://doi.org/10.3389/fpsyg.2016.00354> PMID: [27014160](https://pubmed.ncbi.nlm.nih.gov/27014160/)
65. Reissland N, Francis B, Buttanshaw L, Austen JM, Reid V. Do fetuses move their lips to the sound that they hear? An observational feasibility study on auditory stimulation in the womb. *Pilot Feasibil Stud.* 2016;2(1):1–7.
66. Lagercrantz H, Changeux J-P. The emergence of human consciousness: from fetal to neonatal life. *Pediatr Res.* 2009;65(3):255–60. <https://doi.org/10.1203/PDR.0b013e3181973b0d> PMID: [19092726](https://pubmed.ncbi.nlm.nih.gov/19092726/)
67. Narayanan DZ, Takahashi DY, Kelly LM, Hlavaty SI, Huang J, Ghazanfar AA. Prenatal development of neonatal vocalizations. *eLife.* 2022;11:e78485. <https://doi.org/10.7554/eLife.78485> PMID: [35880740](https://pubmed.ncbi.nlm.nih.gov/35880740/)