

# Neural circuits underlying mother's voice perception predict social communication abilities in children

Daniel A. Abrams<sup>a,1</sup>, Tianwen Chen<sup>a</sup>, Paola Odriozola<sup>a</sup>, Katherine M. Cheng<sup>a</sup>, Amanda E. Baker<sup>a</sup>, Aarthi Padmanabhan<sup>a</sup>, Srikanth Ryali<sup>a</sup>, John Kochalka<sup>a</sup>, Carl Feinstein<sup>a</sup>, and Vinod Menon<sup>a,b,c,1</sup>

<sup>a</sup>Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA 94305; <sup>b</sup>Program in Neuroscience, Stanford University School of Medicine, Stanford, CA 94305; and <sup>c</sup>Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA 94305

Edited by Michael I. Posner, University of Oregon, Eugene, OR, and approved April 1, 2016 (received for review February 24, 2016)

**The human voice is a critical social cue, and listeners are extremely sensitive to the voices in their environment. One of the most salient voices in a child's life is mother's voice: Infants discriminate their mother's voice from the first days of life, and this stimulus is associated with guiding emotional and social function during development. Little is known regarding the functional circuits that are selectively engaged in children by biologically salient voices such as mother's voice or whether this brain activity is related to children's social communication abilities. We used functional MRI to measure brain activity in 24 healthy children (mean age, 10.2 y) while they attended to brief (<1 s) nonsense words produced by their biological mother and two female control voices and explored relationships between speech-evoked neural activity and social function. Compared to female control voices, mother's voice elicited greater activity in primary auditory regions in the midbrain and cortex; voice-selective superior temporal sulcus (STS); the amygdala, which is crucial for processing of affect; nucleus accumbens and orbitofrontal cortex of the reward circuit; anterior insula and cingulate of the salience network; and a subregion of fusiform gyrus associated with face perception. The strength of brain connectivity between voice-selective STS and reward, affective, salience, memory, and face-processing regions during mother's voice perception predicted social communication skills. Our findings provide a novel neurobiological template for investigation of typical social development as well as clinical disorders, such as autism, in which perception of biologically and socially salient voices may be impaired.**

auditory | voice | reward | brain | children

**T**he human voice is a critical social cue for children. Beyond the semantic information contained in speech, this acoustical signal provides a wealth of socially important information. For example, the human voice provides information regarding who is speaking, a highly salient perceptual feature that has been described as an "auditory face" (1). From the earliest stages of development, human listeners are extremely sensitive to the different voices in their environment (2), reflecting the importance of this social cue to human interaction and communication.

Listeners are particularly sensitive to the familiar voices encountered in their everyday environment, and arguably the most salient vocal source in a child's life is mother's voice. Mother's voice is a constant and familiar presence in a child's environment, beginning at a time when these vocal sounds and vibrations are conducted through the intrauterine environment to the fetus' developing auditory pathways (3). Early exposure to mother's voice facilitates recognition of this sound source and establishes it as a preferred stimulus: From the first days of life, children can identify their mother's voice and will actively work to hear this sound source in preference to unfamiliar female voices (2). Throughout development, communicative cues in mother's voice convey critical information to guide behavior (4–6) and learning (7). For example, hearing a recording of one's own mother's voice is a source of emotional comfort for preschoolers during stressful situations, even when the content of the speech is meaningless (5). Furthermore, when school-age females experience a stressful situation, hearing

their mother's voice reduces children's cortisol levels, a biomarker of stress, and increases oxytocin levels, a hormone associated with social bonding (4). These studies have highlighted the profound influence that mother's voice has on children's cognitive, emotional, and social function.

Despite the behavioral importance of mother's voice for critical aspects of emotional and social development, little is known about the mechanisms by which socially salient vocal sources shape the developing brain. Near-infrared spectroscopy (8) and EEG (9) studies examining responses to mother's voice have focused on young children ( $\leq 6$  mo old) and have found increased neural activity for mother's voice compared to female control voices; however, the methods used in these studies are unable to provide detailed information about the brain areas and functional circuits underlying the perception of mother's voice. Therefore, a critical question remains: What are the neural representations of a biologically salient vocal source in a child's brain?

To investigate this question, we used functional MRI (fMRI) and measured brain activity in 24 typically developing children (7–12 y old; see [Tables S1](#) and [S2](#)) in response to their mother's voice, an example of a highly socially salient vocal source in a child's life. An important component of our experimental protocol included vocal recording sessions of each participant's mother and two female control voices, both of whom are also mothers and were not known to the study participants, for subsequent presentation during functional brain imaging (Fig. 1A; see [Methods](#) and [Audio Files S1–S6](#) for audio examples). During the recording sessions, mothers produced three four-syllable nonsense words, which were

## Significance

**The human voice provides a wealth of social information, including who is speaking. A salient voice in a child's life is mother's voice, which guides social function during development. Here we identify brain circuits that are selectively engaged in children by their mother's voice and show that this brain activity predicts social communication abilities. Nonsense words produced by mother activate multiple brain systems, including reward, emotion, and face-processing centers, reflecting how widely mother's voice is broadcast throughout a child's brain. Importantly, this activity provides a neural fingerprint of children's social communication abilities. This approach provides a template for investigating social function in clinical disorders, e.g., autism, in which perception of biologically salient voices may be impaired.**

Author contributions: D.A.A. and V.M. designed research; D.A.A., P.O., K.M.C., A.E.B., and A.P. performed research; D.A.A., T.C., A.E.B., S.R., and V.M. contributed new reagents/analytic tools; D.A.A., T.C., P.O., A.E.B., and J.K. analyzed data; and D.A.A., C.F., and V.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

<sup>1</sup>To whom correspondence may be addressed. Email: daa@stanford.edu or menon@stanford.edu.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1602948113/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1602948113/-DCSupplemental).

used to avoid activating semantic systems in the brain (10), thereby enabling a focus on the neural responses to each speaker's vocal characteristics.

We had two primary goals for the data analysis. First, we wanted to probe neural representations and circuits elicited by mother's voice across all participants. We hypothesized that the critical role of mother's voice in social and emotional learning and its function as a rewarding stimulus would facilitate a distinct representation of this sound source in the minds of children, reflected by neural activity and connectivity patterns in auditory, voice-selective (11), reward (12), and social cognition (13) systems in the brain. The second goal of the analysis was to explore individual differences in brain responses to mother's voice among children. We reasoned that children's social communication and language function could potentially account for individual differences in brain responses to mother's voice. Although it is established that children show a range of cognitive and language abilities, it also has been shown that they demonstrate a range of social abilities (14). Given the important contribution of mother's voice to social communication (4–6), we hypothesized that the strength of functional connectivity between voice-selective cortex and reward and affective processing regions would predict social function in neurotypical children.

## Results

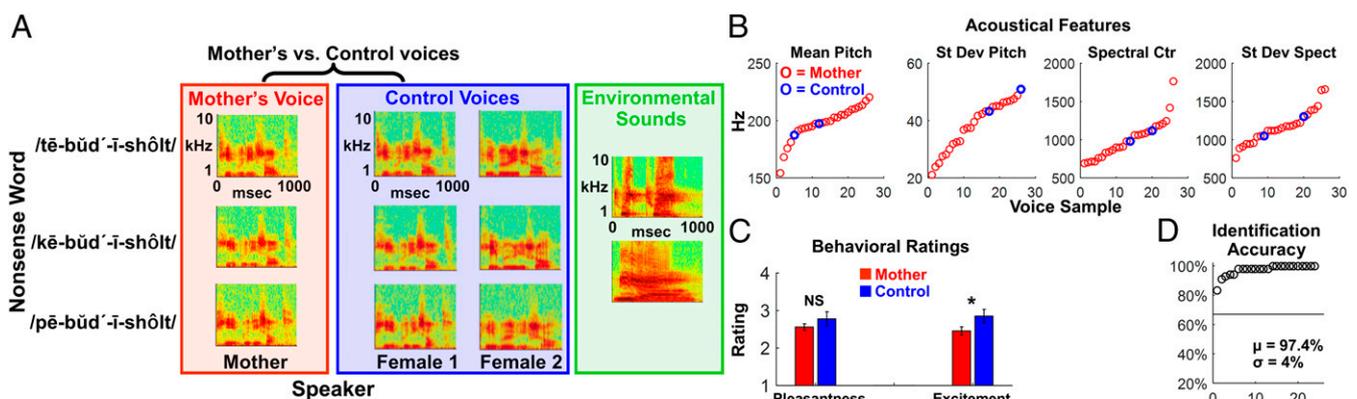
**Acoustical and Behavioral Analysis of Mother's Voice and Control Voices.** We conducted acoustical analyses and behavioral experiments to characterize the physical and perceptual attributes of mother's voice and female control voice samples. The goal of these analyses was to determine if there were differences between mother's voice and female control voice samples that could account for differences in fMRI activity beyond the biological salience of mother's voice. Human voices are differentiated according to a number of acoustical characteristics, including features that reflect the anatomy of the speaker's vocal tract, such as the pitch and harmonics of speech, and learned aspects of speech production, which include speech rhythm, rate, and emphasis (15, 16). Acoustical analysis of the vocal samples used in the fMRI scan showed that control voice samples were qualitatively similar to mother's voice samples across multiple spectrotemporal acoustical features (Fig. 1B).

We next examined perceptual attributes of the stimuli. Of particular interest are the attributes associated with the pleasantness and excitement (a child-friendly proxy for "engagingness") of the

vocal samples: If the vocal characteristics of the mother's voice samples are more rewarding and exciting than those of the female control voices, this difference could potentially account for brain effects associated with hearing mother's voice. We administered a separate behavioral experiment in an independent cohort (i.e., children who did not participate in the fMRI study) of 27 elementary school children (mean age: 11.1 y). In this experiment, participants rated the 24 mother's voice stimuli used in the fMRI experiment and the two female control stimuli based on how pleasant and exciting these voices sounded (*SI Methods*). We found no statistical difference between pleasantness ratings for the control voices and the mean pleasantness ratings for the mother's voice samples (Fig. 1C, *Left*); however, female control voices showed greater excitement ratings than the mother's voice samples ( $P = 0.023$ ) (Fig. 1C, *Right*). Importantly, these behavioral results show that the vocal qualities of the two female control voices used in the fMRI experiment were equally as pleasant as, and were not less exciting than, the mother's voice stimuli.

**Identification of Mother's Voice.** To examine whether children who participated in the fMRI study could identify their mother's voice accurately in the brief vocal samples used in the fMRI experiment, participants performed a mother's voice identification task (*SI Methods*). We found that children identified their mother's voice with a high degree of accuracy (mean accuracy >97%) (Fig. 1D), indicating that brief (<1 s) pseudoword speech samples are sufficient for the consistent and accurate identification of mother's voice.

**Brain Responses to Mother's Voice Compared to Female Control Voices.** In the fMRI analysis, we first identified brain regions that showed greater activation in response to mother's voice compared to female control voices. By subtracting out brain activation associated with hearing female control voices producing the same nonsense words (i.e., controlling for low-level acoustical features, phoneme and word-level analysis, auditory attention, and other factors), we estimated brain responses unique to hearing the maternal voice. We found that mother's voice elicited greater activity in a number of brain systems, encompassing regions important for auditory, voice-selective, reward, social, and visual functions. First, mother's voice elicited greater activation in primary auditory regions, including bilateral inferior colliculus



**Fig. 1.** fMRI experimental design, acoustical analysis, and behavioral results. (A) Randomized, rapid event-related design: During fMRI data collection, three auditory nonsense words, produced by three different speakers, were presented to the child participants at a comfortable listening level. The three speakers consisted of the child's mother and two female control voices. Nonspeech environmental sounds were also presented to enable baseline comparisons for the speech contrasts of interest. All auditory stimuli were 956 ms in duration and were equated for rms amplitude. (B) Acoustical analyses show that vocal samples produced by the participants' mothers were similar to the female control voice samples for individual acoustical measures. (C) Results from behavioral ratings, collected in an independent cohort of children who did not participate in the fMRI study, show that female control voice samples were rated equally as pleasant as, and more exciting than, the mother's voice samples.  $*P < 0.05$ ; NS, not significant. (D) Children who participated in the fMRI study were able to identify their mother's voice with high levels of accuracy, supporting the sensitivity of these young listeners to their mother's voice. The horizontal line represents chance level for the mother's voice identification task.

(IC), the primary midbrain nucleus of the ascending auditory system, and bilateral posteromedial Heschl's gyrus (HG), which contains the primary auditory cortex (Fig. 2). The auditory association cortex of the superior temporal plane, including bilateral planum temporale and planum polare, also showed significantly greater activation in response to mother's voice, with slightly greater activation in the right hemisphere. Next, mother's voice elicited enhanced bilateral activation in voice-selective superior temporal gyrus (STG) and superior temporal sulcus (STS), extending from posterior ( $y = -48$ ) to anterior ( $y = 14$ ) aspects of the lateral temporal cortex. Mother's voice also elicited greater activity in the medial temporal lobe, including the left-hemisphere amygdala, a key node of the affective processing system. Structures of the mesolimbic reward pathway also showed greater activation in response to mother's voice than to female control voices, including the bilateral nucleus accumbens (NAc) and the ventral putamen of the ventral striatum, orbitofrontal cortex (OFC), and ventromedial prefrontal cortex (vmPFC). Mother's voice also elicited greater activation in posterior medial cortex bilaterally encompassing the precuneus and posterior cingulate cortex, a key node of the default mode network (17), which is a system involved in processing self-referential information (18). Additionally, mother's voice elicited increased activity in multiple regions of the occipital cortex, including right-hemisphere intercalcarine, lingual, and fusiform cortex, including overlap with the FG2 subregion of the fusiform, which is associated with visual face processing (19). Greater activation also was evident in the anterior insula (AI) and the dorsal anterior cingulate cortex (dACC), two key structures of the salience network (20). Finally, preference for mother's voice was evident in frontoparietal regions, including right-hemisphere pars opercularis [Brodmann area (BA) 44] and triangularis (BA 45), and in bilateral angular, supramarginal, and precentral gyri. The signal level in the majority of these brain regions showed increased activity relative to baseline in response to mother's voice (see *SI Methods* and *Figs. S1–S4* for results from signal-level analysis). No brain regions showed significantly greater activation for female control voices compared to mother's voice.

We explored sources of variance in participants' voxelwise responses by performing whole-brain covariate analyses using social and language scores as covariates. Results from whole-brain analysis showed that standardized measures of social or language abilities did not show significant correlations with brain activity levels in reward, affective, or salience-processing regions.

**Brain Responses to Female Control Voices Compared to Nonvocal Environmental Sounds.** We next examined whether the extensive brain activation in response to mother's voice (Fig. 2) is specific to this stimulus or, alternatively, if a similar extent of activation is elicited by female control voices when compared to nonvocal environmental sounds. This particular comparison was used in a seminal study examining the cortical basis of vocal processing in adult listeners (11), and results from the current child sample are

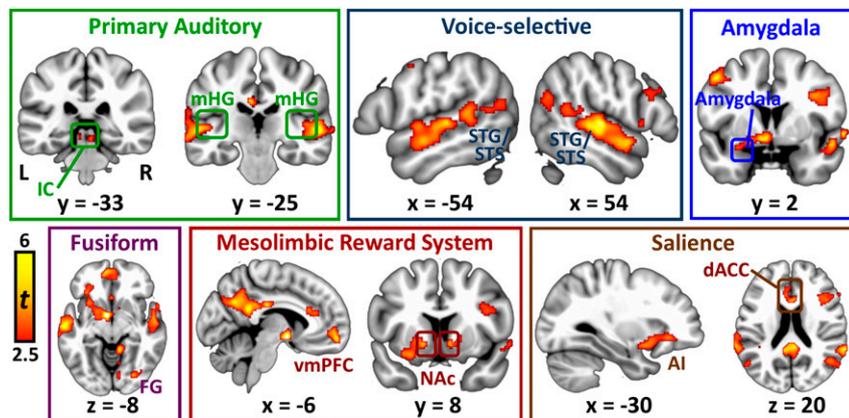
consistent with this previous work, showing strong activation in bilateral voice-selective STG and STS (Fig. S5) for this contrast. Moreover, female control voices elicit activity in bilateral amygdala and supramarginal gyri and in left-hemisphere medial HG (mHG). Importantly, this analysis comparing female control voices and environmental sounds failed to identify reward, salience, and face-processing regions or the IC. Together, these results not only demonstrate that responses to mother's voice are highly distributed throughout a number of brain systems but also show that activity in many of these regions, encompassing reward, salience, and face-processing systems, is specific to mother's voice.

**Analysis of Control Voices.** We next examined whether the presence of pleasant vocal features in the control voices could elicit increased activity in brain systems activated by mother's voice (Fig. 2). This analysis was based on independent behavioral ratings of the vocal stimuli, which revealed that vocal pleasantness ratings were significantly greater for one of the female control voices compared to the other control voice ( $P < 0.001$ ). Both whole-brain and region of interest (ROI) analyses showed no differences in brain response between the two control voices in auditory, voice-selective, face-processing, reward, salience, or default mode brain regions (see *SI Methods, Control voice analysis*). These results indicate that more intrinsically pleasant vocal characteristics alone are not sufficient to drive brain activity in the wide range of brain systems engaged by mother's voice.

**Functional Connectivity During Mother's Voice Processing.** The brain regions identified by the voxelwise analysis of mother's voice identified multiple functional systems encompassing primary auditory and voice-selective temporal cortex, cortical structures of the visual ventral stream, and heteromodal regions associated with affective and reward function and salience detection. A prominent hypothesis states that the STS is a key node of the speech perception network that connects low-level auditory regions with heteromodal regions important for reward and affective processing of these sounds (21). Therefore, our next analysis examined the functional connectivity of the STS, using the generalized psychophysiological interaction (gPPI) model, with the goal of identifying the brain network that shows greater connectivity during mother's voice compared to female control voice perception.

Given the broad anterior–posterior expanse of STS/STG that showed greater activity for mother's voice compared to female control voices (Fig. 2), we placed gPPI seeds bilaterally in posterior, mid, and anterior STG/STS (see *Table S3* for seed coordinates). Surprisingly, group results did not reveal significant brain connectivity during mother's voice perception between any of the STS/STG seeds and affective and reward processing regions or structures of the salience network and visual ventral stream.

**Individual Differences in Functional Connectivity During Mother's Voice Processing.** We then investigated individual differences in



**Fig. 2.** Brain activity in response to mother's voice. Compared to female control voices, mother's voice elicits greater activity in auditory brain structures in the midbrain and superior temporal cortex (*Upper Left*), including the bilateral IC and primary auditory cortex (mHG) and a wide extent of voice-selective STG (*Upper Center*) and STS. Mother's voice also elicited greater activity in occipital cortex, including fusiform gyrus (FG) (*Lower Left*), and in heteromodal brain regions serving affective functions, anchored in the amygdala (*Upper Right*), core structures of the mesolimbic reward system, including NAc, OFC, and vmPFC (*Lower Center*), and structures of the salience network, including the AI and dACC (*Lower Right*). No voxels showed greater activity in response to female control voices compared to mother's voice.



responses for mother's voice? Sensory representations are sharpened and strengthened for behaviorally salient stimuli (28, 29), ostensibly to facilitate their rapid identification, and it is plausible that the behavioral importance of mother contributes to the strengthening of sensory representation for her voice in auditory regions in her child's brain. A potential mechanism for the enhancement of auditory cortical responses is the coincident activity of auditory and reward circuitry: Previous work has shown that stimulation of dopaminergic neurons in reward circuitry during auditory stimulus presentation selectively enhances auditory cortical representations for the presented sounds (30). We hypothesize that the identification of mother's voice as a rewarding stimulus drives synchronous activity in auditory and reward circuitry and facilitates the strengthening of mother's voice representations throughout auditory cortex.

Our results also show, for the first time to our knowledge, that mother's voice drives neural activity in a number of key nodes of the reward circuit (12), including the NAc, OFC, and vmPFC. Activity in this circuit reflects both the anticipation and the experience of preferred stimuli, including music (31, 32), whose rewarding nature has received considerable attention (33). Vocal sounds, on the other hand, are not typically considered a "rewarding" category of sounds, possibly because of their ubiquity in everyday life, and structures of the reward circuit are not considered part of the canonical speech-perception network (27). During development, however, mother's voice is thought to constitute a rewarding stimulus to young children (34), and this initial attraction to the sounds of speech is thought to guide early language acquisition (35, 36). Our findings suggest that the rewarding nature of mother's voice can be detected even in late childhood and demonstrate that brief samples of salient speech stimuli have preferred access to the distributed reward circuit. More generally, we propose that the reward circuit plays an active role in multiple aspects of speech perception, including identifying preferred speech sources and positively valenced emotional cues provided by personally relevant voices.

Our findings further identify a strong link between children's social communication abilities—their ability to interact and relate with others—and speech-based brain connectivity. Specifically, our results show that functional connectivity between voice-selective STS and the NAc of the reward circuit, the amygdala, salience network, FG, and hippocampus, a key structure for memory function, predicts social communication abilities. This result is consistent with previous findings that intrinsic connectivity between voice-selective STS and reward structures and the amygdala predicts social communication abilities in children with autism spectrum disorders (37). Results from the current study advance our understanding of individual differences in neurotypical children through the use of distinct and biologically salient speech stimuli. Surprisingly, despite prominent individual differences related to social communication, voice-selective STS did not show significantly greater connectivity for mother's voice than for female control voices at the group-averaged level. These results suggest that tightly coordinated neural activity between voice-selective STS and brain regions serving reward and affective processes is specific to children with greater social communication abilities.

An important question is whether brain responses to mother's voice simply reflect the intrinsic pleasantness of this vocal source compared to control voices. We addressed this question using several additional analyses. First, we behaviorally characterized all vocal stimuli and found that female control voice samples were rated equally as pleasant as the mother's voice samples. Second, we found that, despite female control samples being equally pleasant, mother's voice elicited greater activity and connectivity compared to control voices in auditory, voice-selective, face-processing, reward, salience, and default mode network regions; in contrast, no brain areas showed greater engagement to control voices compared to mother's voice. Third, analysis of the two control voices, which had shown significantly different pleasantness ratings, revealed comparable brain responses across these key brain systems. Together, these results

indicate that vocal pleasantness is not sufficient to drive brain activity in the wide range of brain systems engaged by mother's voice.

Another question is whether brain responses to mother's voice simply reflect a familiarity response to a recognizable vocal source (38, 39). A number of distinguishing features of the current results suggest that mother's voice elicits a more specialized form of response than the response identified in these previous findings. For example, familiarity effects in previous studies have failed to identify primary auditory cortex, structures of the reward network, including the NAc, OFC, and vmPFC, or key nodes of the salience network, including the dACC and AI (20). Moreover, if familiarity were the only variable driving responses to mother's voice, one would not expect to see a strong relation between children's social skills and brain connectivity during mother's voice processing. Based on these findings, we hypothesize that brain responses to mother's voice reflect specialized representations of a salient source for social learning in a child's life.

In conclusion, we have identified key functional systems and circuits underlying the perception of a foundational sound source for social communication in a child: mother's voice. Critically, the degree of engagement of these functional systems represents a biological signature of individual differences in social communication abilities. Our findings provide a novel neurobiological template for the investigation of normal social development as well as clinical disorders such as autism (37), in which perception of biologically salient voices may be impaired (40).

## Methods

**Participants.** The Stanford University Institutional Review Board approved the study protocol. Parental consent and children's assent were obtained for all evaluation procedures, and children were paid for their participation in the study. All children were required to have a full-scale intelligence quotient (IQ) >80, as measured by the Wechsler Abbreviated Scale of Intelligence (WASI) (41). Participants were the biological offspring of the mothers whose voices were used in this study (i.e., none of our participants were adopted, and therefore none of the mothers' voices were from an adoptive mother), and all participants were raised in homes that included their mothers. Participants' neuropsychological and language characteristics are provided in [Tables S1 and S2](#), respectively. Details are provided in [SI Methods](#).

**Stimuli.** Stimuli consisted of the three nonsense words, "teebudishawlt," "keebudishawlt," and "peebudishawlt," produced by the participant's mother and by two female control voices produced by women who are also mothers (Fig. 1; see [Audio Files S1–S6](#) for audio examples). A second class of stimuli included in the study was nonspeech environmental sounds. Details are provided in [SI Methods](#).

**Data Acquisition Parameters.** All fMRI data were acquired in a single session at the Richard M. Lucas Center for Imaging at Stanford University. Functional images were acquired on a 3-T Signa scanner (General Electric) using a custom-built head coil. Details are provided in [SI Methods](#).

**fMRI Task.** Auditory stimuli were presented in 10 separate runs, each lasting 4 min. The order of stimulus presentation was the same for each subject. Details are provided in [SI Methods](#).

**fMRI Preprocessing.** Details of fMRI preprocessing are provided in [SI Methods](#).

**Voxelwise Analysis of fMRI Activation.** The goal of the voxelwise analysis of fMRI activation was to identify brain regions that showed differential activity levels in response to mother's voice, female control voices, and environmental sounds. Details are provided in [SI Methods](#).

**Effective Connectivity Analysis.** Effective connectivity analysis was performed using gPPI (42), a method more sensitive than psychophysiological interaction (PPI) to context-dependent differences in connectivity. Details are provided in [SI Methods](#).

**Brain-Behavior Analysis.** Regression analysis was used to examine the relationship between brain signatures of mother's voice perception and social and language skills. Social function was assessed using the Social Communication subscale of the SRS-2 (22). For our measure of language

function, we used the CELF-4 (23), a standard instrument for measuring language function in neurotypical children. Regression analyses were conducted using the Core Language Score of the CELF, a measure of general language ability. Brain-behavior relationships were examined using analysis of both activation levels and effective connectivity. Details are provided in *SI Methods*.

**Functional Brain Connectivity and Prediction of Social Function.** To examine the robustness and reliability of brain connectivity between STS and reward, affective, salience detection, and face-processing brain regions for predicting social communication scores, we performed a confirmatory cross-validation (CV) analysis that employs a machine-learning approach with balanced fourfold CV combined with linear regression (25). Details are provided in *SI Methods*.

1. Belin P, Fecteau S, Bédard C (2004) Thinking the voice: Neural correlates of voice perception. *Trends Cogn Sci* 8(3):129–135.
2. DeCasper AJ, Fifer WP (1980) Of human bonding: Newborns prefer their mothers' voices. *Science* 208(4448):1174–1176.
3. Kisilevsky BS, Hains SM (2011) Onset and maturation of fetal heart rate response to the mother's voice over late gestation. *Dev Sci* 14(2):214–223.
4. Seltzer LJ, Prosski AR, Ziegler TE, Pollak SD (2012) Instant messages vs. speech: Hormones and why we still need to hear each other. *Evol Hum Behav* 33(1):42–45.
5. Adams RE, Passman RH (1979) Effects of visual and auditory aspects of mothers and strangers on the play and exploration of children. *Dev Psychol* 15(3):269–274.
6. Mumme DL, Fernald A, Herrera C (1996) Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Dev* 67(6):3219–3237.
7. Liu HM, Kuhl PK, Tsao FM (2003) An association between mothers' speech clarity and infants' speech discrimination skills. *Dev Sci* 6(3):F1–F10.
8. Imafuku M, Hakuno Y, Uchida-Ota M, Yamamoto J, Minagawa Y (2014) "Mom called me!" Behavioral and prefrontal responses of infants to self-names spoken by their mothers. *Neuroimage* 103:476–484.
9. Purhonen M, Kilpeläinen-Lees R, Valkonen-Korhonen M, Karhu J, Lehtonen J (2004) Cerebral processing of mother's voice compared to unfamiliar voice in 4-month-old infants. *Int J Psychophysiol* 52(3):257–266.
10. Binder JR, Desai RH, Graves WW, Conant LL (2009) Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb Cortex* 19(12):2767–2796.
11. Belin P, Zatorre RJ, Lafaille P, Ahad P, Pike B (2000) Voice-selective areas in human auditory cortex. *Nature* 403(6767):309–312.
12. Haber SN, Knutson B (2010) The reward circuit: Linking primate anatomy and human imaging. *Neuropsychopharmacology* 35(1):4–26.
13. Adolphs R, Tranel D, Damasio AR (1998) The human amygdala in social judgment. *Nature* 393(6684):470–474.
14. Constantino JN, Todd RD (2003) Autistic traits in the general population: A twin study. *Arch Gen Psychiatry* 60(5):524–530.
15. Bricker PD, Pruzansky S (1976) Speaker recognition. *Contemporary Issues in Experimental Phonetics*, ed Lass NJ (Academic, New York), pp 295–326.
16. Hecker MH (1971) Speaker recognition. An interpretive survey of the literature. *ASHA Monogr* 16:1–103.
17. Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proc Natl Acad Sci USA* 100(1):253–258.
18. Gusnard DA, Akbudak E, Shulman GL, Raichle ME (2001) Medial prefrontal cortex and self-referential mental activity: Relation to a default mode of brain function. *Proc Natl Acad Sci USA* 98(7):4259–4264.
19. Caspers J, et al. (2014) Functional characterization and differential coactivation patterns of two cytoarchitectonic visual areas on the human posterior fusiform gyrus. *Hum Brain Mapp* 35(6):2754–2767.
20. Menon V, Uddin LQ (2010) Saliency, switching, attention and control: A network model of insula function. *Brain Struct Funct* 214(5-6):655–667.
21. Belin P, Bestelmeyer PE, Latinus M, Watson R (2011) Understanding voice perception. *Br J Psychol* 102(4):711–725.
22. Constantino JN, Gruber CP (2012) *Social Responsiveness Scale, Second Edition (SRS-2)* (Western Psychological Services, Torrance, CA).
23. Semel E, Wiig EH, Secord WH (2003) *Clinical Evaluation of Language Fundamentals* (Psychological Corporation, San Antonio, TX), 4th Ed.
24. Evans TM, et al. (2015) Brain structural integrity and intrinsic functional connectivity forecast 6 year longitudinal growth in children's numerical abilities. *J Neurosci* 35(33):11743–11750.
25. Cohen JR, et al. (2010) Decoding developmental differences and individual variability in response inhibition through predictive analyses across individuals. *Front Hum Neurosci* 4:47.
26. Supekar K, et al. (2013) Neural predictors of individual differences in response to math tutoring in primary-grade school children. *Proc Natl Acad Sci USA* 110(20):8230–8235.
27. Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8(5):393–402.
28. Wang X, Merzenich MM, Sameshima K, Jenkins WM (1995) Remodelling of hand representation in adult cortex determined by timing of tactile stimulation. *Nature* 378(6552):71–75.
29. Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci* 13(1):87–103.
30. Bao S, Chan VT, Merzenich MM (2001) Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412(6842):79–83.
31. Salimpoor VN, et al. (2013) Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science* 340(6129):216–219.
32. Menon V, Levitin DJ (2005) The rewards of music listening: Response and physiological connectivity of the mesolimbic system. *Neuroimage* 28(1):175–184.
33. Huron D (2006) *Sweet Anticipation: Music and the Psychology of Expectation* (MIT Press, Cambridge, MA).
34. Lamb ME (1981) Developing trust and perceived effectiveness in infancy. *Advances in infancy research*, ed Lipsitt LP (Ablex, Norwood, NJ), Vol 1, pp 101–127.
35. Curtin S, Vouloumanos A (2013) Speech preference is associated with autistic-like behavior in 18-month-olds at risk for autism spectrum disorder. *J Autism Dev Disord* 43(9):2114–2120.
36. Vouloumanos A, Curtin S (2014) Foundational tuning: How infants' attention to speech predicts language development. *Cogn Sci* 38(8):1675–1686.
37. Abrams DA, et al. (2013) Underconnectivity between voice-selective cortex and reward circuitry in children with autism. *Proc Natl Acad Sci USA* 110(29):12060–12065.
38. Shah NJ, et al. (2001) The neural correlates of person familiarity. A functional magnetic resonance imaging study with clinical implications. *Brain* 124(Pt 4):804–815.
39. von Kriegstein K, Kleinschmidt A, Sterzer P, Giraud AL (2005) Interaction of face and voice areas during speaker recognition. *J Cogn Neurosci* 17(3):367–376.
40. Uddin LQ, et al. (2013) Saliency network-based classification and prediction of symptom severity in children with autism. *JAMA Psychiatry* 70(8):869–879.
41. Wechsler D (1999) *Wechsler Abbreviated Scale of Intelligence* (Harcourt, San Antonio, TX).
42. McLaren DG, Ries ML, Xu G, Johnson SC (2012) A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *Neuroimage* 61(4):1277–1286.
43. Glover GH, Law CS (2001) Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magn Reson Med* 46(3):515–522.
44. Abrams DA, et al. (2011) Decoding temporal structure in music and speech relies on shared brain resources but elicits different fine-scale spatial patterns. *Cereb Cortex* 21(7):1507–1518.
45. Abrams DA, et al. (2013) Multivariate activation and connectivity patterns discriminate speech intelligibility in Wernicke's, Broca's, and Geschwind's areas. *Cereb Cortex* 23(7):1703–1714.
46. Luculano T, et al. (2014) Brain organization underlying superior mathematical abilities in children with autism. *Biol Psychiatry* 75(3):223–230.
47. Cox RW (1996) AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29(3):162–173.
48. Bullmore E, et al. (1996) Statistical methods of estimation and inference for functional MR image analysis. *Magn Reson Med* 35(2):261–277.
49. Forman SD, et al. (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): Use of a cluster-size threshold. *Magn Reson Med* 33(5):636–647.
50. Ward BD (2000) *Simultaneous Inference for fMRI Data*. AFNI 3dDeconvolve Documentation (Medical College of Wisconsin, Milwaukee, WI).
51. Smith SM, et al. (2004) Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23(Suppl 1):S208–S219.
52. Vul E, Harris C, Winkielman P, Pashler H (2009) Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition. *Perspect Psychol Sci* 4(3):274–290.
53. Mühlau M, et al. (2006) Structural brain changes in tinnitus. *Cereb Cortex* 16(9):1283–1288.
54. Abrams DA, et al. (2013) Inter-subject synchronization of brain responses during natural music listening. *Eur J Neurosci* 37(9):1458–1469.
55. Morosan P, et al. (2001) Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage* 13(4):684–701.
56. Abrams DA, Nicol T, Zecker S, Kraus N (2008) Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J Neurosci* 28(15):3958–3965.
57. Gustafson K, House D (2001) *Fun or boring? A Web-Based Evaluation of Expressive Synthesis for Children* (Eurospeech, Aalborg, Denmark).

Please see *SI Methods* for (i) *Movement Criteria for Inclusion in fMRI Analysis*, (ii) *Signal-Level Analysis*, (iii) *Stimulus Design Considerations*, (iv) *Stimulus Recording*, (v) *Stimulus Postprocessing*, (vi) *Pleasantness and Excitement Ratings for Vocal Stimuli*, and (vii) *Postscan Speaker Identity Recognition Task*, and *SI Results* for (i) *fMRI Sex Difference Analysis* and (ii) *Control Voice Analysis*.

**ACKNOWLEDGMENTS.** We thank all the children and their parents who participated in our study, E. Adair for assistance with data collection, the staff at the Lucas Center for Imaging for assistance with data collection, and H. Abrams and C. Anderson for help with stimulus production. This work was supported by NIH Grants K01 MH102428 (to D.A.A.), K25 HD074652 (to S.R.), and DC011095 and MH084164 (to V.M.) and by the Singer Foundation and the Simons Foundation (V.M.).