Human infants' preference for speech decomposed: Meta-analytic evidence

Cécile Issard[a,1], **Sho Tsuji**[b], **and Alejandrina Cristia**[a]

[a]Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Études Cognitives, École Normale Supérieure, 29, rue d'Ulm, 75005 Paris, France.; [b]International Research Center for Neurointelligence, The University of Tokyo, Tokyo, Japan.

This manuscript was compiled on August 12, 2020

Previous experimental studies suggested that infants show a preference for speech over other sounds, but the mechanism underlying the preference remains unclear, with some results pointing to a preference for (a) familiar sounds; (b) vocal sounds; or (c) natural sounds more generally, with potential changes over development. To shed light on whether some of these mechanisms may be at play, we conducted a meta-analysis of experiments testing speech preference in infants. Synthesizing data from 1071 infants across 34 experiments, we found that infants reliably preferred speech over other sounds, but this preference was not significantly moderated by familiarity, vocal quality, or naturalness of the competitor. Also, we found no effect of age: infants showed the same strength of preference throughout the first year of life. Speech therefore appears to be preferred from birth, even to other natural or vocal sounds. These results nuance current views of the development of speech perception, and call for further investigation of the phenomenon.

Infants | Meta-analysis | Conspecific | Speech perception | Developmental tuning

In gregarious species, the ability of individuals to orient their attention towards conspecifics' vocalizations is crucial for social interactions and communication. Correspondingly, numerous results have demonstrated processing advantages for conspecific vocalizations in a wide variety of species, including primates and birds (1). Previous work has suggested that in humans, communication skills develop from early infancy, manifesting in an early preference for speech over other types of sound (2–4). This preference may emerge through a variety of simpler mechanisms, saliently including a preference for familiar over unfamiliar sounds, natural over artificial sounds, and vocal over non-vocal sounds (Figure 1). Here, we synthesize the available empirical data on infants' preferences for speech over non-speech sounds to assess the size of this effect, and the explanatory role of these mechanisms.

One mechanism could involve familiarity: Perhaps infants prefer speech to other sounds because speech is a frequent sound in their experience. Newborns prefer their native speech to prosodically distinct foreign speech (e.g. 5, 6), which supports a preference for sound patterns heard frequently in the womb. If this mechanism is at play in speech preference, infants should show a stronger preference for speech over other sounds when tested in their native language, and a weaker preference when tested with a foreign language.

A second mechanism could involve a preference for natural over artificial sounds. Natural sounds are those produced by biological systems, such as heart beats, step sounds, or animal vocalizations, and environmental/geophysical sounds, such as wind, rain, or the sound of a river. Natural sounds are processed more efficiently by the auditory system, from the cochlea (7) to the auditory cortex (see 8, for a review). If this mechanism (partially) explains speech preferences, then we should observe that the preference for speech over artificial competitors is stronger than the preference observed when contrasting speech against another natural sounds.

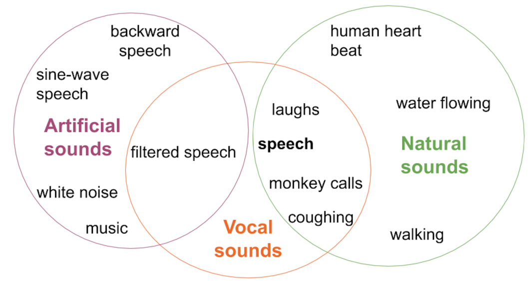

Fig. 1. Speech is a natural, vocal sound.

A third mechanism would rely on a preference for vocal sounds, which may even be shared with other non-human primates given the prevalence of oral communication across primate species and birds. Vocal sounds have acoustic signatures, since they are characterized by modulations introduced by the vocal tract, with harmonically related energy peaks. If this mechanism explains at least partially speech preferences, the preference for speech over non-vocal competitors is stronger than the preference observed when contrasting speech against another vocal sounds.

Development may affect the preference for speech in various

Significance Statement

Numerous species, such as primates and songbirds, use vocal communication to build social interactions. Accordingly, they must be able to pay attention to their conspecifics' vocalizations. In humans, this ability would be rooted in an initial preference for vocal sounds, that would tune to speech during the first year of life. We leverage meta-analysis to examine this theory and shed light on the mechanisms that may explain this preference. We show that speech is preferred by infants readily from birth, even to other natural or vocal sounds, independently of the language spoken, and that this preference is stable across the first year of life. Our findings support a parallel with the development of face perception, which may serve as a foundation for social cognition.

Cécile Issard collected the meta-analytic data, with input from Alejandrina Cristia. Cécile Issard, Sho Tsuji, and Alejandrina Cristia analyzed data. Cécile Issard and Alejandrina Cristia wrote the manuscript. Cécile Issard, Sho Tsuji and Alejandrina Cristia reviewed the manuscript.

The authors declare no competing interests.

[1]To whom correspondence should be addressed. E-mail: cecile.issardgmail.com

ways. For example, whereas newborns do not prefer speech over monkey calls, there is some evidence that three-month-olds do (3). More broadly, both maturation and experience could affect the extent to which naturalness, familiarity, and vocal quality affect infants' preferences as a function of age.

A meta-analytic approach. In sum, previous results on infants' preference for speech are broadly compatible with a preference for natural over artificial, vocal over non-vocal, and/or familiar over unfamiliar sounds, potentially interacting with infants' age. In this paper, we seek to directly test these potential mechanisms by employing a meta-analytic approach.

Meta-analyses are quantitative syntheses of experiments testing a comparable phenomenon with comparable approaches. In a nutshell, meta-analyses allow us to draw statistics over multiple experiments. This is important because any single experiment is a noisy window onto the underlying reality of a given phenomenon: Despite experimentalists' best efforts, a given experiment can at best tell us about what a specific group of participants, presented with a specific set of stimuli, in a specific point in time has done. As a result, even if we assume that all experiments that are carried out are systematically published, meta-analyses are useful because they allow us to integrate those individual and specific results into a larger picture. By integrating data across different laboratories, they provide evidence for the generalizability of effects across labs, and facilitate comparisons between experimental results.

Moreover, meta-analyses allow direct comparison and integration of data in a principled statistical approach. To begin with, data drawn from larger samples and with higher precision is given more weight than smaller and less precise estimates. By combining studies to obtain larger samples, meta-analyses can reveal small effects not obvious in individual experiments.

Additionally, a meta-analysis allows us to statistically test and/or statistically control for different factors. This allows us to integrate data from experiments that vary in their methodology, as well as test the effect of factors that are not part of the original design, by redescribing the stimuli used as a function of those factors. For instance, a study measuring preference for native speech over native backward speech provides data on a natural versus artificial contrast, as well as a vocal versus non-vocal contrast.

We can also draw a developmental timeline across the age range covered by the literature, beyond age groups tested within papers. This is particularly useful for potential changes with age, as in the field of infant studies it is common to make age-related statements using the demonstrably problematic method of concluding that there is an interaction without actually testing for it statistically (9). It is therefore important to directly test for interactions with an actual statistical approach. To give an example from a previous developmental meta-analysis, it had been proposed that infants' preference for novel or familiar items related to infants' age such that, all things equal, younger infants showed familiarity preferences whereas older infants exhibited novelty preferences (10). However, **(author?)** (11) found stable familiarity preferences for word segmentation in natural speech across the first two years; and **(author?)** (12) found a stable novelty effect for artificial grammars implemented in synthesized speech, whereas those implemented in natural speech led to stable familiarity preferences. Meta-analyses are therefore important to statistically and systematically test the theoretical predictions proposed in

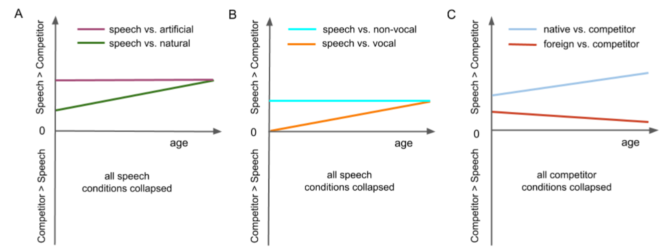

Fig. 2. Hypothesized pattern of preference: the x axis shows age, the y axis represents the effect size derived from the contrast between a speech condition and a competitor condition (preference for speech over the competitor is plotted up; the lower quadrants are empty because we do not predict a preference for the competitor over speech). A: Speech contrasted to natural (green) or artificial (purple) competitors. B: Speech contrasted to vocal (orange) or non-vocal (cyan) competitors. C: Collapsing across competitors, separating speech in a foreign language (red); speech in the native language (blue).

qualitative reviews, and refine current readings of a literature.

Finally, meta-analyses offer tools to detect publication bias in the literature. By aggregating all the available evidence for a phenomenon, we can see if the distribution of effect sizes has an unexpected shape, typically with an excess of positive results due to the difficulty to publish null or negative results. We can further integrate this information, and derive a new estimate of the overall effect size.

The present study. Although we were primarily motivated by a theoretical quest, we know meta-analyses provide a unique vantage point on a body of work as a whole. We therefore first check for how strong infants' preference for speech over other types of sounds is according to the public body of literature. We additionally assess this body of data for evidence of publication bias.

We then turn to our key interest, namely shedding light on the potential mechanisms underlying infants' speech preferences. Meta-regressions assess whether the proposed mechanisms of familiarity, naturalness, and vocal quality drive this preference, and how the preference develops over the first year of life. Assuming all three mechanisms are at play, and further assuming that the definition of the preferred stimulus narrows with age, we predicted that infants will show (see Figure 2):

1. a greater preference for native speech over non-speech as a function of age, but a smaller preference for foreign speech over non-speech with age; 2. a greater preference for speech over other natural sounds as a function of age, but a preference for speech over artificial sounds that is stable over development; 3. a greater preference for speech over other vocal sounds as a function of age, but a stable preference for speech over non-vocal sounds over development.

Results

Average effect size. We found a total of 18 publications (labeled with an asterisk in the reference list) reporting 34 experiments, for a total of 1071 infants from 0 to 12 months (1.5 to 380.5 days), and 54 effect sizes, see Figure 3 (See Supplementary Information S1 for a complete description of the database). We integrated all effect sizes in a meta-analytic regression without any moderator, and found an average effect size g of 0.31 (SE = 0.06, t = 5.10, CI = [0.19 , 0.44]) (Supplementary Figure S2), corresponding to a medium effect size.

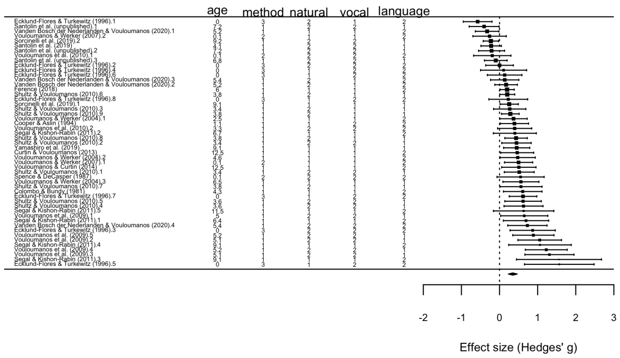

Fig. 3. Forest plot of effect sizes available in the literature, along with their respective moderator status.

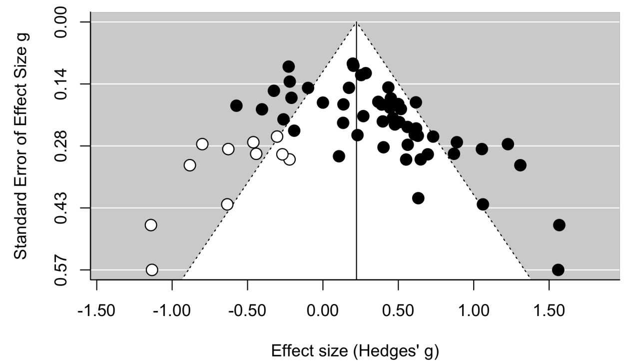

Fig. 4. Funnel plot of effect sizes and their respective standard errors. Black dots: effect sizes observed in the literature. White dots: missing effect sizes, suggestive of a publication bias. Vertical line: average effect size after filling the missing effect sizes.

Heterogeneity among effect sizes was estimated at $\tau^2 = 0.12$ ($I^2 = 76.80\%$), which was significant (Q = 190.35, p<0.01) despite the removal of outliers before running the model. This strongly suggests differences across experiments, and invites analyses using moderators.

Publication bias. We checked for the presence of a potential publication bias in the body of literature by studying the relationship between standard errors of effect sizes as a function of Hedges' g (Figure 4). A regression test on these data was significant (z = 7.55, p < 0.01), as was the Kendall's tau rank correlation test for funnel plot asymmetry (Kendall's tau = 0.48, p < 0.01), consistent with a publication bias in the literature.

To check whether this bias fully explains infants' speech preference, we symmetrized the funnel plot with the "trim and fill" method (13). To symmetrize the funnel plot, 11 (SE = 4.71) missing experiments were needed on the left side of the plot. The corrected effect size was estimated at 0.22 (SE = 0.06) after filling in the 11 missing experiments, which is still significantly different from zero. Thus, even correcting for a potential publication bias, we still find statistical evidence for infants' preferring speech over competitors.

Moderator analyses. We then tested if heterogeneity could be accounted for by the mechanistic explanations described in our introduction. Following our hypotheses, we fit a meta-analytic model with the following moderators: (1) mean age of children; (2) familiarity with the language used (native or foreign); (3) naturalness of the contrastive sound (coded as yes if it was natural and no otherwise); (4) vocal quality of the contrastive sound (coded as yes if it was vocal and no otherwise).

Given the number of effect sizes available in the literature, these moderators were specified without interactions to avoid overfitting. None of the moderators was significant (Table 1). We also tested each of our three hypotheses by three separate models for each moderator and its interaction with age. None of them yielded a significant effect or interaction (Figures 6, 7, 5; supplementary results S5).

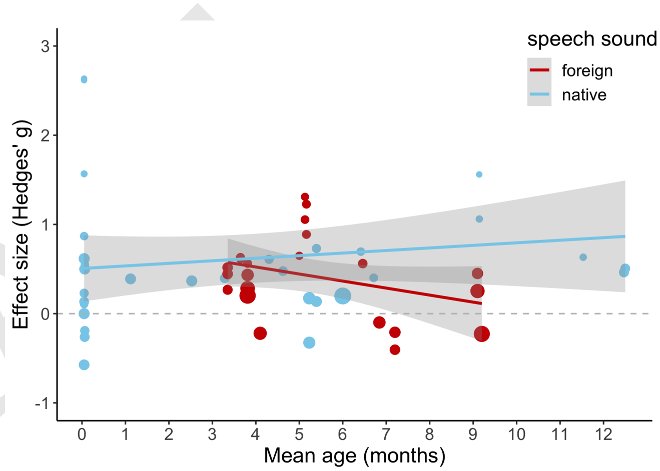

Fig. 5. Effect sizes as a function of age and familiarity with the speech sounds. The size of each dot is inversely proportional to the variance. Positive effect sizes reflect a preference for the speech sound, negative effect sizes reflect a preference for the competitor sound.

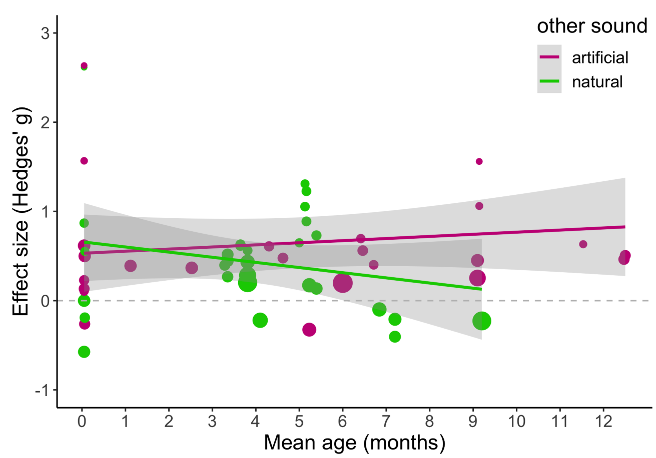

Fig. 6. Effects sizes as a function of age and natural quality of the competitor. The size of each dot is inversely proportional to the variance. Positive effect sizes reflect a preference for the speech sound, negative effect sizes reflect a preference for the competitor sound.

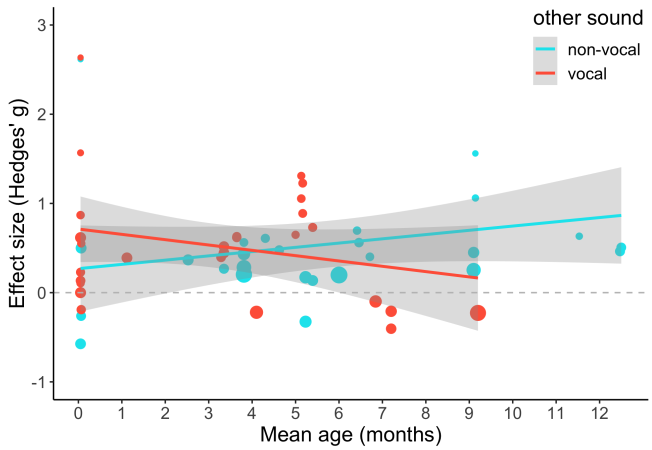

Fig. 7. Effect sizes as a function of age and vocal quality of the competitor. The size of each dot is inversely proportional to the variance. Positive effect sizes reflect a preference for the speech sound, negative effect sizes reflect a preference for the competitor sound.

Discussion

Our meta-analysis synthesizes the available literature on infants' preference for speech sounds. When all experiments were considered together with no moderators, we found a sizable intercept (g=0.31). Our meta-analysis shows that this preferential processing of speech sounds is observable from birth on. We had hypothesized age to play a major role, because it may correlate with a reshaping of the category definition for speech itself. Indeed, experiments comparing processing of human speech against human non-speech as well as animal vocalizations more generally (3, 14) often discuss these age-related differences in categorization of these sounds. Surprisingly, age did not significantly moderate the overall preference for speech, as shown by the null estimate of this moderator (Table 1), nor did it interacted with any other moderators. This result was replicated in a separate model for age only, which showed a significant intercept, similar to the intercept found in the meta-regression with no moderator (Supplementary results S5). Crucially, we left age uncentered. This intercept therefore provides an estimate of the effect size at age 0, i.e. at birth. Moreover, the scatterplot of effect sizes as a function of age reveals clearly no change with age even when plotted without other moderators (Supplementary figure S6). This null effect of age, combined with the sizable intercept, confirms that infants reliably prefer speech over other types of sounds from birth.

The significant heterogeneity we found among the literature suggests that underlying factors modulate this effect. We had predicted infants' speech preference to be larger when the competitor was an artificial sound than when it was a natural one; when the competitor was non-vocal; and when the speech was in the infants' native language. In fact, we were unable to disprove the null hypothesis of no difference for all three factors. Distributions of effect sizes for experiments varying along the three dimensions widely overlap, and the confidence intervals of all conditions overlap almost exactly (Figures 6, 7, and 5). Moreover, Table 1 shows that the estimate for all these factors is close to zero (the maximum being 0.21). Our findings therefore suggest that none of these parameters fully explain infants' preference for speech sounds.

In other words, from birth on, infants show a preference for speech, which cannot be accounted by the three explanations tested here: naturalness, vocalness, or familiarity with the native language. It is possible that infants prefer speech because of its complex acoustic structure and fast transitions (15). Spectral or temporal modulations taken separately are not sufficient to elicit neural responses similar to the ones elicited by speech (16). However, speech is characterized by joint spectrotemporal modulations at specific rates (17). It is possible that infants are sensitive to this specific spectro-temporal structure (though see 18, showing that they only explain neural responses in primary auditory cortex, suggesting that other factors contribute to the behavioral response in later processing stages). Testing this explanation would require to compute the modulation spectra of the actual stimuli used in the experiments. Thus, we recommend interested researchers to deposit their stimuli in a public archive such as the Open Science Framework (19).

A capacity to preferentially listen to speech sounds from birth suggests that infants are born with the capacity to recognize their conspecifics' communication signals. This parallels what has been proposed for faces: Infants are born with the capacity to orient their attention toward them, even without any prior exposure to faces (20, 21). This would stem from basic perceptual abilities present at birth, namely that the visual system would be tuned to a spatial structure that correspond to those of faces (20, 21). As newborns have never been exposed to such visual stimuli before, it would reflect general properties (i.e. filters) of the visual system. Similarly, the auditory system could be tuned to a spectro-temporal structure that speech presents. The combination of this non-specific bias with the systematic variations of the auditory environment would result in preferential responses to speech from birth. However, contrary to faces, fetuses are exposed to speech that is low-pass filtered by the womb throughout the last trimester of gestation (22, 23). It is therefore possible that prenatal experience with low-pass filtered speech helps infants to form a representation of speech, by refining the response properties of the auditory system to speech. The fact that familiarity with the language used in the experiment did not modulate infants' preference suggests that this effect is not triggered by familiarity with the sounds of the native language. Infants would therefore form a representation that is specific enough to discriminate speech from other natural or vocal sounds, but general enough to be independent of the language spoken.

Ultimately, preferential processing of speech may support higher level cognitive tasks. The human species is a gregarious one. Detecting speech signals would allow to integrate it with other sensory percepts, such as faces, to form multisensory representations of conspecifics (24). This would lay the track for social cognition. Identifying speech signals and paying attention to them would allow infants to form complex representations of the sensory world, that they can manipulate cognitively. Consistently, infants could categorize visual stimuli (i.e., associate a label to a category of objects) when they were associated to speech, but not pure tones or backward speech (25–27). Interestingly, infants categorized visual stimuli when presented with speech, melodies, monkey (28), or lemur vocalizations (27). These results support the idea that infants may preferentially process complex sounds. Finally, the preference itself may also be a meaningful index of processing that can be used to identify children at risk (29). Understanding this phenomenon is therefore crucial for both theoretical and clinical advances.

We want to close by drawing the reader's attention to methodological insights the present study provides. First, the median sample size at present is 16, which is close to the field standard (30), but much lower than current recommendations (31). Well-chosen sample sizes are crucial for powerful experiments (32). Our meta-analysis provides the average effect size across the literature, which will, in turn, allow researchers to run power analyses to determine the sample size they need in future experiments. Second, the distribution of effect sizes in the literature is consistent with publication bias, in view of a strong asymmetry of the funnel plot. The trim-and-fill method suggested 11 points may be missing. Given that we have 54 effect sizes in total, a fifth more would be missing. The missing experiments are in the negative section, i.e., a preference *against* speech. The literature being biased toward positive effect sizes, the true effect size might be smaller than the average one we found with the available literature (see vertical line on Figure 4 for an estimate of the true effect size). To correct this bias, we invite researchers to use registered reports, in which manuscript are submitted before data are collected (33). Our dataset can be community-augmented, and we invite researchers investigating this phenomenon to complement it with any data they would have (), whatever the results and publication status.

Our meta-analysis revealed uneven distributions of experiments across age and stimulus dimensions. One may think that this uneven age distribution could be one potential reason for the lack of an age effect. However, by aggregating the numerous individual studies of a literature, meta-analyses gain statistical power. As such, meta-analytic results have more cumulative explanatory value than any single study. Previous experiments suggested that infants initially perceive speech and other vocal sounds as a common category, and then narrow it to speech around five months of age. With a sample size of 1071 infants, covering a much wider age range than individual experiments, our meta-analysis provide evidence for an alternative view of the phenomenon, namely that infant prefer speech from birth onwards, even to other natural or vocal sounds, and that this effect is not driven by familiarity with the sounds of the language used. This clearly points to the value of meta-analysis: to advance a field and inspire follow-up studies. In particular, future experiments should test infants older than 9 months. Language production gains in complexity at about this age (34), which could affect infants' speech preference. Experiments using natural vocal stimuli as competitor, and foreign speech as target, would contribute to fill in the gap in the litterature that our meta-analysis revealed.

SI Datasets. Our data is fully available in the corresponding OSF repository: http://tidy.ws/bqjc4U

Materials and Methods

This meta-analysis was carried out following PRISMA recommendations (35). We provide information on all steps (including PRISMA checklist, data, and code) for full transparency and accountability via online supplementary materials.

Literature search. We composed the initial list of papers with suggestions by experts (authors of this work); one google scholar search

Table 1. Statistical results of meta-regression with all moderators. The intercept corresponds to the effect size when speech is in a foreign language, and the competitor is natural, and vocal, at age 0. The moderator estimates correspond to changes in the intercept when the target stimuli are in the native language (familiarity); the competitor is artificial (naturalness); and the competitor is non-vocal (vocal quality).

	estimate	SE	t	95% C.I.
intercept	0.27	0.14	1.93	-0.03 - 0.57
familiarity	-0.03	0.12	-0.22	-0.27 - 0.22
naturalness	0.21	0.11	1.97	-0.02 - 0.44
vocal quality	-0.09	0.14	-0.66	-0.4 - 0.21
age	<0.001	<0.001	0.07	<0.001 - <0.001

(*("speech preference" OR "own-species vocalization") AND infant - "infant-directed"*), the same search in PubMed and PsycInfo (last searched on 2019-09-24); and a google alert. We also inspected the reference lists of all included papers. Finally, we emailed a major mailing list to ask for missing data. We received two replies, one of which revealed a formerly undiscovered published study, and communicated unpublished data (36).

Inclusion criteria. After a first screening based on titles and abstracts using more liberal inclusion criteria, we decided on final inclusion based on full paper reading. We included experiments that tested human infants from birth to one year of age, and contrasted speech sounds with any other type of sound, measuring behavioral preferences to the sounds (e.g., looking times). If a paper reported results from neurotypical and at-risk infants, we included only the data from the neurotypical group.

Given our key interest in the preference for speech over other sounds, we excluded studies that contrasted two different speech sounds (e.g., foreign vs. native language, or adult vs. child-directed speech, or mother vs. stranger's voice); or two different non-speech sounds (e.g., backward speech vs. animal vocalizations). In addition, we excluded experiments where the contrast presented to the infants could not be coded according to our three mechanistic explanations. This meant the exclusion of experiments where speech was presented in the mother's voice (which thus confounds between speech and individual voice recognition for our familiarity factor). Finally, we excluded neuroimaging experiments to avoid mixing results from different brain regions with different response profiles. We included published (i.e., journal articles) as well as unpublished works (i.e., doctoral dissertations) as long as sufficient information was provided.

A PRISMA flow chart summarizes the literature review and selection process (Supplementary Figure S2). The full list of the papers that were inspected together final inclusion decisions are available in a decision spreadsheet (see the online supplementary materials).

Coding. Data were coded by the first author. In addition, 20% of the papers were randomly selected to be coded by the last author independently, with disagreements resolved by discussion. There were 10 disagreements out of a total of 260 fields filled in, and they were indicative of the coders not following the codebook, which led to a revision of all data in four variables.

The critical variables for our purpose are key stimuli characteristics, infant age, and testing method (central fixation, high amplitude sucking, head-turn preference procedure). As for key stimuli characteristics, we coded familiarity, naturalness, and vocal quality, as follows.

For familiarity, we considered the language in which the speech sounds were recorded (native or foreign).

For naturalness, the competitor sound was coded as natural if it was produced by a biological organism without any further acoustic manipulation. Natural competitors included animal calls, environmental sounds (e.g. wind or water sounds), heartbeat, bird song, non-speech vocalizations (e.g. laughters or coughs). If the authors applied acoustic manipulations, the competitor was coded as artificial. Artificial competitors included sine-wave speech, filtered speech, white noise, instrumental music, and speech with altered rhythmic structure. In the case of filtered speech, the modulations introduced by the vocal tract are still present at the retained frequencies, and formant transitions are consistent with vocal production constraints. For this reason, filtered speech can be considered as vocal but not natural. The only exception was for newborn experiments presenting low-pass filtered speech mimicking the filtering applied by the womb. Because the womb acts as a low-pass filter, fetuses are exposed to low-pass filtered speech as a natural stimulus. Given the recency of the intra-uterine environment to newborns (about 2 days), we coded these as natural.

For vocal quality, the competitor sound was considered as vocal if it was produced by an animal vocal tract (human or not), either original or modified. Vocal competitors included non-speech vocalizations, animal calls, bird songs, and filtered speech. Non-vocal competitors included backward speech (that has abrupt closures that cannot be produced by the vocal tract), white-noise, environmental sounds, instrumental music, heartbeat, and sine-wave speech (that lacks the harmonic structure introduced by the natural resonance of the vocal tract).

We coded all the statistical information reported in the included papers. If reported, we coded the mean score and the standard deviation for speech, and the other sound separately. When infant-level data was provided, we recomputed the respective mean scores and standard deviations based on the reported individual scores. If reported, we also coded the t-statistic between the two sound conditions, or an F-statistic provided this was a two-way comparison. If effect sizes were directly reported as a Cohen's d or a Hedges' g, we also coded this.

Effect sizes. Once the data were coded, we extracted effect sizes, along with their respective variance. Effect sizes were standardized differences (Cohen's d) between response to speech vs. the competitor. If effect sizes were not directly reported in the papers, we computed them using the respective means and SDs (37), or a t- or F-statistic (38). As our effect sizes came from within-subject comparisons (e.g., looking time of the same infant during speech and during monkey calls), we needed to take into account the correlation between the two measurements in effect sizes and effect size variances computations. We computed this correlation based on the t-statistic, the respective means and SDs (37) if they were all reported; or imputed this correlation randomly if not. We finally calculated the variance of each effect size (37). Cohen's d were transformed to Hedges' g by multiplying d by a correction for small sample sizes based on the degree of freedom (39).

We did not center age because our hypotheses included a developmental progression from birth to the end of the first year of life. We were therefore interested in the intercept at age 0 (i.e., birth).

Analyses use the R (40) package Robumeta (41), which allows us to fit meta-analytic regressions that take into account the correlated structure of the data, when repeated measures are obtained from the same infant groups within papers.

ACKNOWLEDGMENTS. This work was supported by an Agence Nationale de la Recherche grant to A.C. (ANR-17-CE28-0007 LangAge, ANR-16-DATA-0004 ACLEW, ANR-14-CE30-0003 MechELex, ANR-17-EURE-0017); and the J. S. McDonnell Foundation Understanding Human Cognition Scholar Award to A.C. We thank Chiara Santolin, Martin Zettersten, Jenny Saffran, Christina Vanden Bosch Der Nederlanden, and Athena Vouloumanos for contributing unpublished data.

1. A Poremba, J Bigelow, B Rossi, Processing of communication sounds: Contributions of learning, memory, and experience. *Hear. Res.* **305**, 31–44 (2013).
2. A Vouloumanos, JF Werker, Listening to language at birth: evidence for a bias for speech in neonates. *Dev. science* **10**, 159–164 (2007).
3. A Vouloumanos, MD Hauser, JF Werker, A Martin, The tuning of human neonates' preference for speech. *Child development* **81**, 517–527 (2010).
4. L Ecklund-Flores, G Turkewitz, Asymmetric headturning to speech and nonspeech in human newborns. *Dev. Psychobiol.* **29**, 205–217 (1996).
5. J Mehler, et al., A precursor of language acquisition in young infants. *Cognition* **29**, 143–178 (1988).
6. C Moon, RP Cooper, WP Fifer, Two-day-olds prefer their native language. *Infant Behav. Dev.* **16**, 495–500 (1993).
7. EC Smith, MS Lewicki, Efficient auditory coding. *Nature* **439**, 978–982 (2006).
8. A Mizrahi, A Shalev, I Nelken, Single neuron and population coding of natural sounds in auditory cortex. *Curr. Opin. Neurobiol.* **24**, 103–110 (2014).
9. A Gelman, H Stern, The Difference Between "Significant" and "Not Significant" is not Itself Statistically Significant. *The Am. Stat.* **60**, 328–331 (2006).
10. MA Hunter, EW Ames, A multifactor model of infant preferences for novel and familiar stimuli. *Adv. Infancy Res.* **5**, 69–95 (1988).
11. C Bergmann, A Cristia, Development of infants' segmentation of words from native speech: a meta-analytic approach. *Dev. Sci.* **19**, 901–917 (2016).
12. A Black, C Bergmann, Quantifying infants' statistical word segmentation: A meta-analysis. (Cognitive Science Society), pp. 124–129 (2017).
13. S Duval, R Tweedie, Trim and Fill: A Simple Funnel-Plot–Based Method of Testing and Adjusting for Publication Bias in Meta-Analysis. *Biometrics* **56**, 455–463 (2000).
14. NM McDonald, et al., Infant brain responses to social sounds: A longitudinal functional near-infrared spectroscopy study. *Dev. Cogn. Neurosci.* **36**, 100638 (2019).
15. S Rosen, P Iverson, Constructing adequate non-speech analogues: what is special about speech anyway? *Dev. Sci.* **10**, 165–168 (2007).
16. Y Minagawa-Kawai, A Cristià, I Vendelin, D Cabrol, E Dupoux, Assessing Signal-Driven Mechanisms in Neonates: Brain Responses to Temporally and Spectrally Different Sounds. *Front. Psychol.* **2** (2011).
17. NC Singh, FE Theunissen, Modulation spectra of natural sounds and ethological theories of auditory processing. *The J. Acoust. Soc. Am.* **114**, 3394 (2003).
18. SV Norman-Haignere, JH McDermott, Neural responses to natural and model-matched stimuli reveal distinct computations in primary and nonprimary auditory cortex. *PLOS Biol.* **16**, e2005127 (2018).
19. ED Foster, A Deardorff, Open Science Framework (OSF). *J. Med. Libr. Assoc. : JMLA* **105**, 203–206 (2017).
20. J Morton, MH Johnson, CONSPEC and CONLERN: A Two-Process Theory of Infant Face Recognition. *Psychol. Rev.* **98**, 164–181 (1991).
21. C Turati, Why Faces Are Not Special to Newborns: An Alternative Account of the Face Preference. *Curr. Dir. Psychol. Sci.* **13**, 5–8 (2004).
22. D Querleu, X Renard, F Versyp, L Paris-Delrue, G Crèpin, Fetal hearing. *Eur. J. Obstet. & Gynecol. Reproductive Biol.* **28**, 191–212 (1988).
23. JP Lecanuet, C Granier-Deferre, Speech Stimuli in the Fetal Environment in *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*, NATO ASI Series, eds. B de Boysson-Bardies, S de Schonen, P Jusczyk, P McNeilage, J Morton. (Springer Netherlands, Dordrecht), pp. 237–248 (1993).
24. A Vouloumanos, MJ Druhen, MD Hauser, AT Huizink, Five-month-old infants' identification of the sources of vocalizations. *Proc. Natl. Acad. Sci. United States Am.* **106**, 18867–18872 (2009).
25. AL Fulkerson, SR Waxman, Words (but not Tones) facilitate object categorization: Evidence from 6- and 12-month-olds. *Cognition* **105**, 218–228 (2007).
26. AL Ferry, SJ Hespos, SR Waxman, Categorization in 3- and 4-Month-Old Infants: An Advantage of Words Over Tones. *Child Dev.* **81**, 472–479 (2010).
27. AL Ferry, SJ Hespos, SR Waxman, Nonhuman primate vocalizations support categorization in very young human infants. *Proc. Natl. Acad. Sci. United States Am.* **110**, 15231–15235 (2013).
28. AL Fulkerson, RA Haaf, The Influence of Labels, Non-Labeling Sounds, and Source of Auditory Input on 9- and 15-Month-Olds' Object Categorization. *Infancy* **4**, 349–369 (2003).
29. A Sorcinelli, J Ference, S Curtin, A Vouloumanos, Preference for speech in infancy differentially predicts language skills and autism-like behaviors. *J. Exp. Child Psychol.* **178**, 295–316 (2019).
30. C Bergmann, et al., Promoting Replicability in Developmental Research Through Meta-analyses: Insights From Language Acquisition Research. *Child Dev.* **89**, 1996–2009 (2018).
31. LM Oakes, Sample Size, Statistical Power, and False Conclusions in Infant Looking-Time Research. *Infancy* **22**, 436–469 (2017).
32. JP Simmons, LD Nelson, U Simonsohn, False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychol. Sci.* **22**, 1359–1366 (2011).
33. A Kiyonaga, JM Scimeca, Practical Considerations for Navigating Registered Reports. *Trends Neurosci.* **42**, 568–572 (2019).
34. DK Oller, RE Eilers, AR Neal, HK Schwartz, Precursors to speech in infancy: The prediction of speech and language disorders. *J. Commun. Disord.* **32**, 223–245 (1999).
35. D Moher, A Liberati, J Tetzlaff, DG Altman, TP Group, Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine* **6**, e1000097 (2009).
36. C Santolin, M Zettersten, JR Saffran, Infants' preference for non-native speech versus birdsong. (2020) original-date: 2020-07-15T23:50:27Z.
37. MW Lipsey, DB Wilson, *Practical meta-analysis*, Practical meta-analysis. (Sage Publications, Inc, Thousand Oaks, CA, US), (2001).
38. WP Dunlap, JM Cortina, JB Vaslow, MJ Burke, Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol. Methods* **1**, 170–177 (1996).
39. M Borenstein, LV Hedges, JPT Higgins, HR Rothstein, *Introduction to Meta-Analysis.* (John Wiley & Sons), (2011) Google-Books-ID: JQg9jdrq26wC.
40. R Core Team, R: A Language and Environment for Statistical Computing (2018).
41. LV Hedges, E Tipton, MC Johnson, Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Methods* **1**, 39–65 (2010).