What Can We Perceive In Infant Vocalization?

Alanna A. Beyak, Olivia Cadieux, Matt Cook, Carly S. Cressman, Barbie Jain, Jarod A. Joshi, Spenser L. Martin, Michael Mielniczek, Sara Montazeri, Jolyn Sawatzky, Bradley C. Smith, Jackie Spear, Essence I. Perera, Thomas Thompson, Derek Trudel, Jianjie Zeng, and Melanie Soderstrom

Department of Psychology, University of Manitoba

Author Note

Melanie Soderstrom: https://orcid.org/0000-0003-3212-5775

This study was registered with Open Collaborative Science (GUID: 2a6b4).

We have no known conflict of interest to disclose. We would like to thank BabbleCor (https://osf.io/rz4tx/) for allowing us to use their data set for this project.

This study was completed in partial fulfilment of the University of Manitoba course PSYC 4540 T39 (honours undergraduate) and PSYC 7310 T34 (graduate), Open and Collaborative Science: Theory and Practice.

Correspondence concerning this article should be addressed to Melanie Soderstrom, Department of Psychology, P404 Duff Roblin Bldg, 190 Dysart Rd, University of Manitoba, Winnipeg, MB, R3T 2N2, Canada. Email: M_Soderstrom@umanitoba.ca

Abstract

Infant vocalization is a well-studied area of development, however, there is a noticeable gap in the literature regarding adult identification of infant vocalization. Authors of the present study collected data from 626 undergraduate students who listened to 100-500 ms audio clips of infant vocalization. Researchers asked participants to identify infants in the audio clips as male/female, English/non-English, and 0-7 months/8-18 months/19-36 months of age. Participants were unable to determine the sex of the infant better than chance but were able to determine the infant's language and age significantly better than chance, $t(463) = 4.4618$, $p < .001$, and $t(463) = 17.714$, $p < .001$, respectively. Exploratory follow-up analyses did not reveal an effect of caregiving experience, childcare experience, or participant gender on a participants' ability to correctly identify the infant's age or language. This research has implications for determining what is and is not perceivable in infant vocalizations. This is an underrepresented topic in infant research as most work has demonstrated what infants can perceive; not what caregivers can. This is an important contribution because infant language development has been demonstrated to include a complex social dynamic between adults and infants.

Keywords: infant vocalization, babble, BabbleCor, language acquisition, adult identification of infant vocalizations

What Can We Perceive In Infant Vocalization?

Most research in infant language development focuses on the infant perspective. However, infants' communicative experiences do not happen in a vacuum but rather in interactions with adult caregivers. Examining this process from the adult perspective is underrepresented in the literature. The ability of an adult to identify infant characteristics such as age, language, and sex could be informative for our understanding of the role that caregivers play in infant language development. Variables such as the adult listener's gender and infant caregiving experience may also be important factors that predict their ability to communicate with infants and discriminate between infant characteristics.

The interactions between caregivers and infants have shown to be integral in the development of prelinguistic vocalizations. Goldstein and Schwade (2008) conducted a study investigating the effect of social feedback on preverbal vocal development. In this experiment, mothers were instructed to either respond contingently or non-contingently to infant babble. Infant vocalizations were measured before and after caregiver responses. The researchers discovered that infants restructured their vocalizations when their mothers responded directly to their babbling. Infants would modify their babbling to fit the same phonological form as their caregivers. These results revealed that some aspects of vocal development require social feedback from the infant's caregiver.

A study conducted by Warlaumont et al. (2014) similarly tested the role of social feedback on the language development of children with and without autism. Across trials, the researchers found that adults respond more readily to speech-related infant vocalizations than non speech-related vocalizations. Additionally, in typical development, infants were more likely to produce a speech-related vocalization if an adult was sensitive to their last vocalization. Thus, increased adult responsiveness can produce more speech-related vocalizations. This increase in speech-related vocalizations will consequently increase adult responsiveness. This evidence supports the existence of a social feedback loop, further solidifying the role of infant-adult interaction in speech development.

Research examining adults' abilities to determine characteristics of infants based on their vocalizations is dated and somewhat inconsistent. One study by Olney and Scholnick (1976) found that when researchers provided infant vocalizations to adult listeners they were able to determine the relative age of an infant. Olney and Scholnick also found that adults were unable to identify the linguistic community that an infant belonged to. Furthermore, Atkinson et al. (1968) noted that adult listeners struggle to identify infants' native languages. However, de Boysson-Bardies et al. (1984) suggested that adults may have some ability to determine the language that an infant is learning.

These findings suggest that adults may use linguistic cues to identify and discriminate between types of preverbal speech and characteristics of the preverbal infants, though researchers have primarily focused on the former. Also, variables such as the listener's gender and caregiving experience are somewhat absent from the relevant analyses that examine preverbal discriminative ability. Some previous studies, however, have investigated the effects of preverbal discrimination training and child-rearing experience on infant speech indentification. Gladding (1979) found that, in comparison, trained adult listeners were significantly better at identifying infant cry signals than untrained adult listeners. Therefore, previous exposure to these signals (e.g. child-rearing) may mediate improvements in preverbal identification and discrimination.

In a more recent study, Lindová et al. (2015) found that adults with children were significantly better than adults without children at identifying positive and negative preverbal infant vocalizations. Additionally, the researchers found that younger adults were more accurate at discriminating between the six infant vocalizations when compared to the older adults. However, Lindová et al. also found that a listeners' gender was not a significant factor in discriminating between the infant vocalizations.

A more specific line of research for studying infant language development is in the study of babbling. Babbling is the early stage of language development where utterances are sounds without recognizable meaning. It appears to be a universal behaviour of infants that is even present when the infants are only learning some form of sign language (Petittio, 2004). Babbling is typically categorized into canonical or non-canonical babbling. Canonical babbling can be defined as babble containing a well-formed vowel and a consonant (E.g. "ba" or "da;" Lee et al., 2017). Infant vocalizations (excluding vegetative noises, laughing and crying) that do not have this structure are called non-canonical babble. The ratio of vowels to consonants is often used to mark the transition from non-canonical to canonical babble. This ratio has been shown to decline from birth (at the 0–2 month stage the ratio is 4.5:1) to 8 months; during the 4–8 month stage the ratio becomes 2.8:1 (Kent & Murray, 1982).

Characteristics of Infant Vocalization

There is some research that has explored whether there are differences in infant vocalizations between males and females. Sex related biological markers have been connected with verbal development. Estrogen has been positively correlated with verbal development while testosterone has been negatively correlated (Quast et al., 2016; Adani & Cepenac, 2019). As early as 4 weeks old, it becomes evident that infant girls develop ahead of boys in verbal acquisition (Adani & Cepenac, 2019). Adani and Cepenac (2019) found FOXP2 proteins, that contribute to enabling verbal expression, are less present in the left hemisphere of 4-year-old boys than girls. This research supports the possibility that there are developmental differences in infant vocalizations between the sexes.

A more extensive body of research has examined the extent that babble reflects the infant's ambient language. According to Jakobson (1941/1968), it was initially proposed that infant babble reflected biologically driven patterns that were universal across language environments. Early work also supported the idea that adults could not discriminate between the babble from infants acquiring different languages (Atkinson, 1968). However, a conflicting perspective suggested that babbling increasingly mirrors verbal characteristics of the eventual mother tongue during language development: This effect is called "babbling drift" (Brown, 1958). Continuing research provided considerable support for the interactional theory with vowel sounds in babbling mirroring the mother tongue across languages (Alhaidary & Rvachew, 2018; de Boysson-Bardies et al., 1984, 1989).

Prosody, the intonational and rhythmic properties of a language, has been analyzed in infant babbling (DePaolis et al., 2008). Language groups (for example, Indo-European and Turkic) vary in prosody and this variability is reflected in infant vocalization. However, differences in prosody have not been shown to be detectable in the babble of languages within a language group (e.g., French and Italian; DePaolis et al., 2008).

Research regarding age indicators of babbling are inconclusive (Best et al., 2016; Eimas et al., 1973). There has been a longstanding debate among researchers whether biophysical events or perceptual-linguistic experiences drive the development of infant vocalic utterances (Snow, 2006). For example, Snow (2006) showed that infants follow a “u-shape” development in pitch and intonation. Infants of 6 to 8 months demonstrated similar intonation as those of 21 to 23 months. This supports the theory that the babble of infants at 6–8 months old has biological prelinguistic tendencies. Furthermore, Snow (2006) indicated that infants at 18–20 months could actively control pitch, whereas between these two stages researchers observed a regression in intonation and pitch control. Changes in vocal development have also been observed between 0 and 8 months, for example Kent and Murray (1982) demonstrated that structural components of infant babble change significantly during this period. The researchers revealed that the ratio of vowels to consonants in infant speech decline from birth to 8 months.

In addition to pitch and intonation, preverbal infants have demonstrated communicative intent in their speech. Stark et al. (1993) measured the communicative intent of 51 infants aged birth to 18 months, that they divided into four age categories. They found that early sounds (birth–2 months), said to be vegetative sounds, included coughing, burping, crying, grunting and sighing. Infants aged 2–5 months added pleasure sounds like laughter and comfort sounds. During the ages of 6–9 months, the sounds also included objects, such that the acoustic pattern was often altered by an object or hand in the mouth. At 10–18 months, infants could also make sounds with intent of social communication. These sounds included words of desire, rejection, feeling, surprise, greetings, as well as animal noises, car noises, agreement (“yes”) and disagreement (“no”; Stark et al., 1993). Differences in communicative intent may be used as a clear indicator of vocal maturity among infants aged 0-18 months.

Adult Listener Discriminative Ability

Previous literature has examined adults' abilities to recognize and interpret infant vocalizations and compared discrimination ability between males and females (Lindová et al., 2015). Other studies, such as Parson et al. (2017), have investigated the accuracy in discriminating infant vocalizations between individuals with caregiver experience and those without. As exploratory analyses, the current study aims to add to the literature by investigating potential differences between males and females, and between caregivers and non-caregivers in their ability to discriminate between infant vocalizations.

Lindová et al. (2015) examined differences between men and women as well as caregivers and non-caregivers who completed babble discrimination tasks. To do this, they exposed men and women to emotionally loaded infant vocalizations and measured the adults' ability to discriminate between a positively- or negatively-charged vocalization. Although the effects were weak, caregiver experience was positively associated with more accurate discriminations; however, the listener's gender had no significant effect on their accuracy. Lindová et al. also found that age had a small effect on vocalization accuracy, such that younger adults were slightly more accurate compared to older adults. Adult listeners could accurately discriminate between positive and negative emotional vocalizations but had difficulty identifying the specific needs of the infants from the vocalizations. This suggests that adults may require other sensory channels to determine specific infant needs (Lindová et al., 2015).

Another category of caregiver that has shown unique neural and behavioural responses to infant vocalizations is mothers. Parsons et al. (2017) performed a similar study examining mothers and non-mothers' responses to infant vocalizations on a neural level. Using functional magnetic resonance imagining (fMRI), Parsons et al. found that rearing a child (i.e., more experience with an infant) had an impact on neural processing of infant cues compared to non-mothers, suggesting that greater infant experience leads to greater and differential neural reactivity, specifically in the amygdala and the orbitofrontal cortex (OFC). Findings were consistent with previous literature highlighting the positive effect of infant expeirence on vocaliziation discrimination ability (Lindová et al., 2015). Furthermore, Parsons et al. suggested that the increase in neural reactivity represents an increased brain capacity and the shaping of neural responses experienced by parents in the caregiving process. These findings provide evidence to suggest that caregivers and non-caregivers have different brain structures and brain activity, and that experience with infants causes increased activity and shaping of specific neural pathways.

Mother-infant associations during infant vocalization have unique relationships at both the neural (Parsons et al., 2017) and behavioural (Hiraoka et al., 2019) levels, compared to non-mothers. Previous research from Hiaroka et al. (2019) outlined the association of pleasant stimuli to approach behaviour and aversive stimuli to avoidance behaviour. The latest research into the effect of infant vocalizations on maternal approach-avoidance behaviour found the opposite response in mothers (Hiraoka et al., 2019). Mothers reacted immediately with approach behaviours when infants elicited a crying or urgent response, suggesting that responding to infant emotional cues is a process unique to mothers. These results add a novel finding to infant vocalization literature that suggests maternal approach behaviours are a direct behavioural response to infant emotional cues (Hiraoka et al., 2019).

As mentioned before, previous studies have examined specific variables of adult listener discriminative ability such as gender and caregiving experience (Lindova et al., 2015) and mothers versus non-mothers' discriminative ability (Parsons et al., 2017). Taking this into account, our hypotheses are also based on testing the participants ability to identify different infant vocalizations. Contrasting to previous research, this study aims to add to the current literature by examining the preverbal discriminative ability of adult listeners. In the present study, adults listened to audio clips of infant vocalizations and identified the sex, age range, and native langauge of the infants. Furthermore, this study will add to the existing literature by examining the effects of caregiving experience and gender on discriminative ability. Participants will be evaluated on their accuracy to identify specific characteristics of infant vocalizations. In sum, our primary objective will be to examine adults' ability to discriminate the age, language and sex of infant vocalizations. As a second objective, we will examine the role of gender and caregiving experience.

Hypotheses

This study tested three preregistered, confirmatory hypotheses and two exploratory hypotheses. The exploratory hypotheses were contingent on results from the confirmatory hypotheses such that they were conducted on the data examined by the first three hypotheses if the original test produced significant results.

Confirmatory Hypotheses

1. Participants will be able to identify the infant's sex significantly above chance (50%).
2. Participants will be able to identify whether the infant is acquiring English or another language above chance (50%).
3. Participants will be able to identify the infant's age range above chance (33%).

Exploratory Hypotheses

4. Participants with caregiving/childcare experience will be able to identify infant's sex, language and/or age significantly better than other participants.
5. Participants who identify as females will be able to identify infant's sex, language and/or age significantly better than other participants.

Method

Participants

The experiment was initially pilot tested internally for debugging purposes. Further, prior to collecting our experimental sample, data was collected from six participants to confirm that there were no technical issues and that our system to automatically grant participants credit was working correctly. These data were not used for any analysis.

Participants were recruited from the SONA subject pool at the University of Manitoba in Winnipeg, Canada. SONA is an online system that first-year undergraduate psychology students use in order to participate in research studies. Students received one credit toward their introductory psychology class upon completion of the study. A total of 626 students participated in this study. Data from 166 participants were excluded; 87 had significant exposure to/conversational experience with non-English languages represented in the corpus, five identified their gender as “Other”, three reported that they resided in a country other than Canada or the USA, seven gave the same response to all questions for an entire experimental block (delinquent responding), 20 people did not identify English as their first language, 14 failed to respond to the attention check correctly more than five times, and 47 failed to respond to the audio check correctly more than five times (some participants were excluded for multiple reasons). Of the 460 remaining participants, 293 were female and 167 were male. The average length of reported time as a primary caregiver (within the last five years) was 1.15 months (*SD* = 6.49), the average length of time reported as a childcare worker or any area with significant exposure to infants (within the last five years) was 5.04 months (*SD* = 12.34), and the average age of the participants was 19.82 years (*SD* = 3.78). All of the data, including those from pilot testing, are accessible on our GitHub repository: https://github.com/melsod/OCSWinter2020.

Materials

We selected audio clips (100-500 ms) of infant vocalizations collected from the BabbleCor corpus (https://osf.io/rz4tx/; Cychosz et al., 2019). BabbleCor is a public cross-linguistic corpus of infant vocalizations in five different languages. We signed and submitted the Data Sharing Agreement (https://osf.io/m2vbd/) in order to have access to a private meta-data file with information for each clip regarding the infant's sex, age (months) and native language, as well as prior "majority vote" classification of each audio clip (canonical, non-canonical, laughing, crying, and junk), from the BabbleCor's citizen science project.

In selecting the clips used for our experiment, we excluded the Warlaumont corpus where infants' native languages were English and/or Spanish with no way to determine which baby spoke which language. We included all the other corpora but limited our sample to those clips classified as canonical and non-canonical. Our objective was to study the development of language, so we excluded the non-babble clips.

We classified the babies found in the corpus into three age categories (0-7 months, 8-18 months, 19-36 months), two language categories (English, Non-English), and two sex categories (Male, Female). Due to constraints imposed by the corpus we could not get a single sample that would provide sufficient control for all three main hypotheses. That is, we were not able to subset the data in such a way that ensured that equal numbers of English and non-English babies were included, while controlling for equal numbers of male and female babies as well as equal number of babies in each age group. Therefore, we selected two separate samples; one for the sex hypothesis and one for the age and language hypotheses. We had six categories of babies for age and language hypotheses sample (English/non-English+age range groups; e.g, English+0-7) and two categories of babies for the sex hypothesis sample (non-English+8-18+Male/Female).

For the age and language audio samples, we randomly selected two babies from each category and 20 clips from each baby. In total, we had 240 clips for the age and language sample. We used the same process for sex, but only had two categories and selected five female babies and five males. For an added level of control the sex sample was only selected from non-english speaking babies who where 8-18 months old. In total, we had 200 clips for the sex sample. Overall, we selected 440 clips to use in the study. The code to select clips uses R-studio 1.2.5019 and R 3.6.1 (2019-07-05) and is on our github repository: https://github.com/melsod/OCSWinter2020/blob/master/R/pre_experiment/sample_clips_2.R.

Three researchers then manually listened to and made judgements on whether any of the selected clips needed to be excluded. One researcher's judgments were omitted because they were too liberal on what criteria to use to omit clips, that led to the exclusion of about half of the database, however they are still included in our documentation. After excluding the recommended clips we were left with fewer than 10 clips for some of the selected babies. Another researcher then randomly selected and screened unused clips for each of these babies until there were 10 usable clips for each selected baby (one clip added into the age and language sample, three clips added into the sex sample). Then, 10 clips were randomly selected from each baby in our sample. All of the researchers who made judgements on the audio clips were blind to the age, sex and language of the babies or even if two clips belonged to the same baby. This process is documented on our GitHub repository: https://github.com/melsod/OCSWinter2020/blob/master/R/pre_experiment/narrow_sample.R.

In the end, we used 217 final clips from 21 babies (12 babies for age-language and 10 for sex, 1 baby was included in both samples). The selected clips used for testing the age and language hypotheses were identical, while the selected clips used for testing the sex hypothesis were different due to the above-mentioned limitations on counterbalancing in clip selection. After selecting the final clips, we used jsPsych (a JavaScript library for running behavioural experiments in a web browser; https://www.jspsych.org/) version 6.2 to program the experiment and hosted the experiment on GitHub Pages. Finally, we stored the data temporarily using Google Firebase (https://firebase.google.com/), we then downloaded, processed, and analyzed it (all documented on our GitHub Repository: https://github.com/melsod/OCSWinter2020).

Procedure

Data collection took place in March, 2020. Although we planned to collect data in person and in our lab, the entire process was moved online due to the COVID-19 pandemic. Participants connected to our study through the SONA system online portal. They were directed to the experiments URL: https://melsod.github.io/OCSWinter2020/. Participants were instructed to read through our informed consent form that briefly explained the purposes of the experiment and reminded them of their rights. Then, written instructions asked the participants to sit in front of their computer while wearing headphones to answer the experimental questions. Participants were informed that the experiment should take no more than 30 minutes to complete. Participants then answered the questions using a mouse and keyboard. The mean time taken to complete the experiment was 12.69 minutes (1% of observations trimmed to remove extreme outliers). After consenting to participate in the study, participants were presented with instructions on how to complete the study. Within the instructions, there was an attention check where the participants had to type in a random word that was provided in the instructions. Then, participants completed a soundcheck where they heard two words and they were asked to press the key corresponding to the first letter and the last letter of the two words, respectively. If participants needed more than 5 attempts in either the attention check or sound check, they were excluded from the data analysis.

Next, participants were provided with 10 brief (100-500 ms) audio clips of vocalizations from an individual infant. Participants responded to one of the following questions based on the audio clips presented: What is the baby's sex? (Male/Female) What language is spoken in the baby's home? (English/Non-English) Or How many months old is the baby? (0-7 months/8-18 months/19-36 months). Participants were required to listen to all of the audio clips at least once (but were not limited in the number of times they could listen to the audio clips) before responding to the question asked. Participants were asked to select a choice according to their best judgement and their answers were collected and anonymized by the computer. After a response, participants would be presented with another 10 audio clips from a different infant. Experimental questions were asked in a block format (e.g. all the sex trials were presented, followed by all of the language trials, followed by all of the age trials) and the order of the three experimental blocks (six unique conditions) was randomized between participants.

After participants completed the three blocks, they were provided with nine general questions about themselves and their experience with babies:

1. In the last 5 years, estimate how many months you have worked/volunteered in childcare or an area with significant exposure to infants (answer “0” if none)
2. In the last 5 years, estimate how many months you have been the primary caregiver for a child under the age of 3 years (answer “0” if none)
3. What is your age in years?
4. What is your gender? (Male, Female, Other)
5. What is your country of residence? (Canada, USA, Other)
6. Do you have a normal hearing? (Yes, No)
7. Is English your first language? (Yes, No)
8. Do you have significant exposure to/conversational experience with any of the following languages? Select all that apply: (Spanish, Tsimane', Yêlí-Dnye, Tseltal Mayan, Quechua)
9. Do you consider yourself monolingual (only speaking one language)? (Yes, No).

Analysis

We used four exclusion criteria based on participant demographics: (a) if a participant had significant exposure to/conversational experience with non-English languages in the corpus (Spanish, Tsimane', Yêlí-Dnye, Tseltal Mayan, or Quechua); (b) if a participant identified their gender as “Other”; (c) if a participant did not identify English as their first language; or (d) if the participant resided in a country other than Canada or the USA. Eighty-seven participants reported that they had significant exposure to/conversational experience with the non-English languages in the corpus, five participants identified their gender as other, three participants reported that they resided in a country other than Canada or the USA, and 20 participants did not identify that English was their first language. Additionally, we used two exclusion criteria that were based on participant responses: (a) if the participant gave the same response to all of the questions for an entire experimental block (delinquent responding); and (b) if they failed to respond correctly to the attention check more than five times. Seven participants were excluded on the basis of delinquent responding, 14 failed to respond to the attention check properly, and 47 failed to respond to the audio check properly. A response was required for every trial and data was not saved until the end of the experiment so there was no missing data.

We measured the proportion of correct responses for each participant for each of the three questions about the infants' sex, age, and language. Participants' responses were coded in a binary fashion as correct or incorrect. We used separate one-sample proportion t-tests to analyze each of the sex, age, and language research questions individually.

We also conducted exploratory analyses for those hypotheses that were supported by the data. We examined whether the gender, recent childcare experience, and/or caregiving experience of the participants had an effect on the proportion of correct responses in the corresponding hypothesis question set.

All analyses were conducted with R and R-studio (R Core Team, 2018; RStudio Team, 2019). The details of all the data and analysis code can be found on our GitHub (https://github.com/melsod/OCSWinter2020) as well as the OSF page (https://osf.io/2a6b4/) and our preregistrations (https://osf.io/2a6b4/registrations).

Results

Hypothesis 1

As specified in the pre-registered analytic plan, a two-tailed, one-sample proportion t-test was used to determine if participants were able to identify the infant's sex significantly above chance (50%). We found that with a $t(459) = -2.93$, $p = .004$, 95% CI = [0.46, 0.49], there is significant evidence (alpha = 0.01) that people respond differently than chance. However, the direction of the difference was in the opposite direction to what we hypothesized and the effect size is very small ($d = -0.14$). This means that participants are slightly worse than chance at predicting the sex of the infants based on the audio clips presented. See Figure 1 for a histogram of the participants' accuracy in determining the infant's sex.

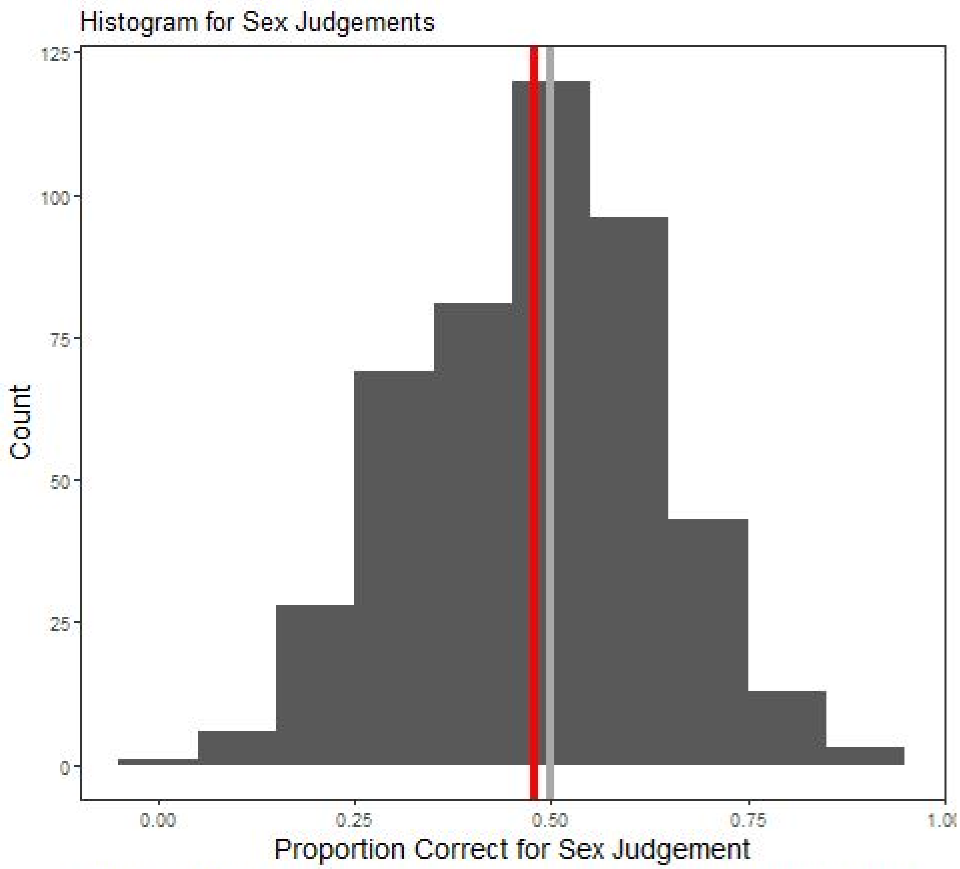

Figure 1. X-axis represents the propotion of correct responses across the 10 experimental trials. Y-axis represents the number of participants with that propotion of correct responses. Mean proportion correct was 0.48 (*SD* = 0.16).

Hypothesis 2

As specified in the pre-registered analytic plan, a two-tailed, one-sample proportion t-test was used to determine if participants were able to identify whether the infant acquired English or a different language above chance (50%). We found that with a $t(459) = 4.94$, $p < .001$, 95% CI = [0.52, 0.55], there is significant evidence (alpha = 0.01) that participants perform better than chance when determining whether the baby was acquiring English. However, this difference corresponds to a small effect size with a Cohen's *d* value of 0.23. See Figure 2 for a histogram of the participants' accuracy in identifying the infant's language.

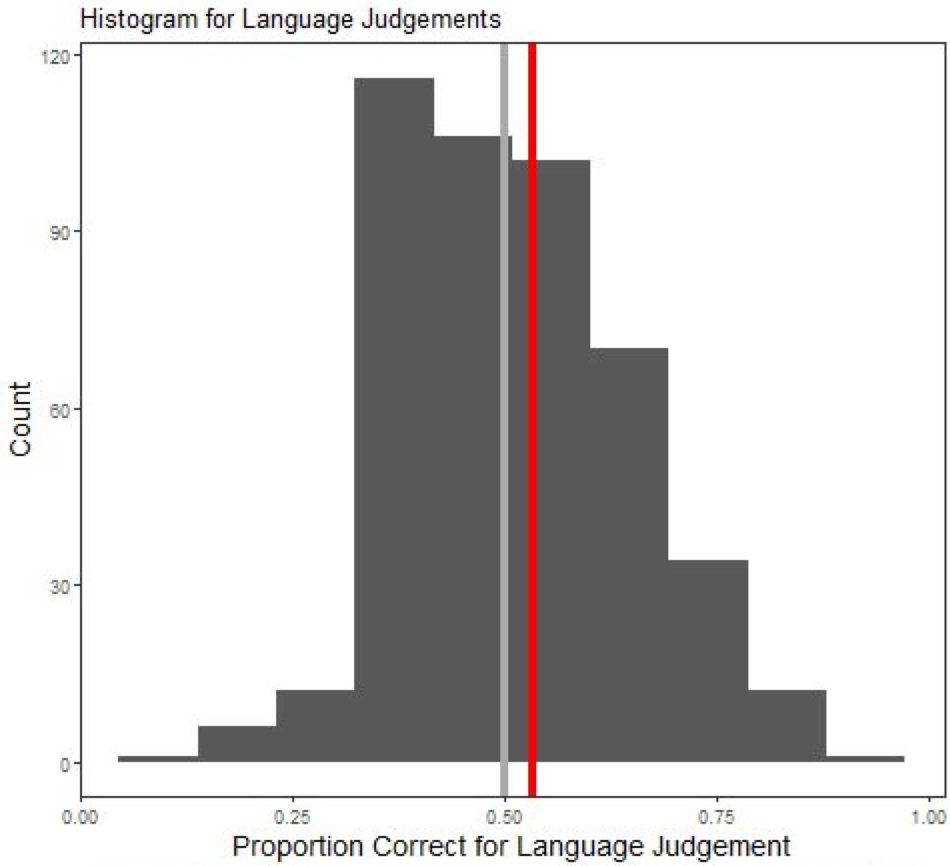

Figure 2. X-axis represents the propotion of correct responses across the 12 experimental trials. Y-axis represents the number of participants with that propotion of correct responses. Mean proportion correct was 0.53 (*SD* = 0.14).

Hypothesis 3

As specified in the pre-registered analytic plan, a two-tailed, one-sample proportion t-test was used to determine if participants were able to identify the infant's age range (from three options) above chance (33%). We found that with a $t(459) = 17.34$, $p < .001$, 95% CI = [0.44, 0.46], there is significant evidence (alpha = 0.01) that participants are able to determine the baby's age better than chance. This difference corresponds to a large effect size with a Cohen's d value of 0.84. See Figure 3 for a histogram of the participants' accuracy in determining the infant's age.

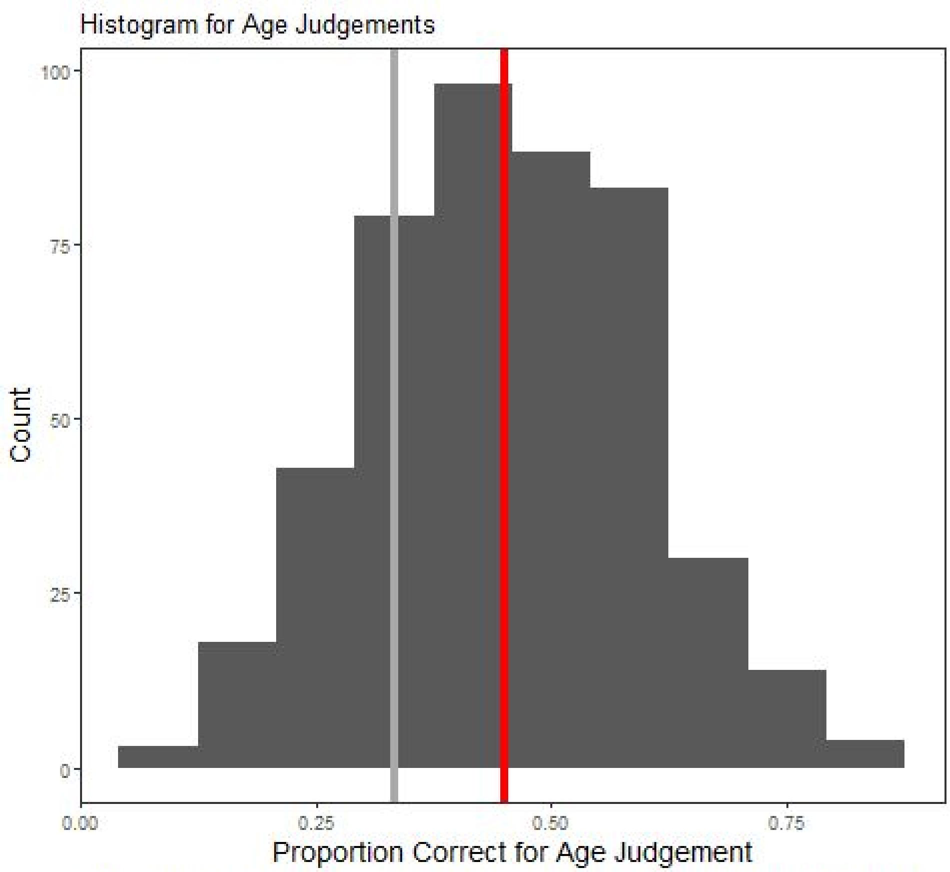

Figure 3. X-axis represents the propotion of correct responses across the 12 experimental trials. Y-axis represents the number of participants with that propotion of correct responses. Mean proportion correct was 0.45 (*SD* = 0.15).

Exploratory Analysis

In order to more fully examine the differences found in our first three hypotheses we conducted planned, but exploratory follow-up analyses for those hypotheses that were supported by the data. To this end we conducted hierarchical within-subjects binary logistic regressions for both the age and language related data. We did not conduct a follow-up analysis on the sex related data because the difference found was in the wrong direction.

AGE

To examine whether gender, recent childcare experience, or caregiving experience of the participant had an effect on identifying infant age, a hierarchical regression analysis was conducted. We first created a baseline model with only an intercept as a random effect. Then recent child interaction experience (recent childcare and caregiving experience separately) were entered to model as a stage one model.

The random intercept model (the baseline model) was a significant predictor of age with $\beta = -0.197$, $z = -7.20$, and with a corresponding $p < .001$. The Akaike information criterion (AIC) value for the baseline model is 7602.9. The stage one model (the child interaction model), contained both recent childcare and caregiving experiences as predictors. We found no significant effect (alpha = 0.01) of childcare experience, with $\beta = 0.005$, a $z = 2.11$, and with a corresponding $p = .035$. There was also no significant effect (alpha = 0.01) for recent caregiving experience with $\beta = -0.007$, a $z = -1.57$, and with a corresponding $p = .118$. The AIC value for the child interaction model was 7602.1. The results indicated that neither childcare nor caregiving experience of participants significantly affects their identification of infants age.

We then compared the baseline model with the child interaction model. The result is chi-square X^2 (2) = 4.84, $p = .089$, suggesting that the effect of recent childcare and caregiving experience together do not improve model fit significantly (alpha = 0.01). However, the AIC value dropped slightly from 7602.9 to 7602.1, suggesting some minor improvement from the baseline model.

To more closely examine whether childcare experience has a significant effect, we conducted an additional (unplanned) model that contained only recent childcare experience as a main variable. The result of the new model showed no effect of childcare experience, $\beta = 0.003$, $z = 1.54$ and $p = .123$. We then compared that new model to the baseline model. The results showed a chi-square value $X^2(1) = 2.37$, $p = .123$. This suggests that childcare experience does not significantly improve the predictive potential of the baseline model. Taken together, these results indicate that neither recent childcare, nor caregiving experience has a significant effect on identifying infant age.

Next, we added gender as another predictive variable into the child interaction model and we called it the gender model (stage two model). Results for this model showed no effect of childcare ($\beta = 0.005$, $z = 2.17$, $p = .030$), caregiving ($\beta = -0.007$, $z = -1.60$, $p = .110$), or gender ($\beta = 0.031$, $z = 0.54$, $p = .587$). The AIC of the gender model was 7603.8.

Finally, we compared the gender model with the child interaction model. The result of comparison was not significant, chi-square value $X^2(1) = 0.29$, $p = .587$, and the AIC value between two models was increasing. These results indicate that adding gender to our model does not significantly improve the predictive potential of the model. Therefore we can conclude that participant gender does not have a significant effect on identifying infant age.

LANGUAGE

To examine whether gender, recent childcare experience, or caregiving experience of the participant had an effect on identifying infant language, a hierarchical regression analysis was conducted. We first created a baseline model with only an intercept as a random effect. Then recent child interaction experience (recent childcare and caregiving experience separately) were entered to model as a stage one model.

The random intercept model (the baseline model) was a significant predictor of language identification with $\beta = 0.129$, $z = 4.79$, and with a corresponding $p < .001$. The AIC value for the baseline model is 7633.4. The stage one model (the child interaction model), contained both recent childcare and caregiving experiences as predictors. We found no significant effect (alpha = 0.01) of childcare experience, with $\beta = .005$, a $z = 1.98$, and with a corresponding $p = .048$. There was also no significant effect (alpha = 0.01) for recent caregiving experience with $\beta = -0.004$, a $z = -0.96$, and with a corresponding $p = .336$. The AIC value for the child interaction model was 7633.4. The results indicated that neither childcare nor caregiving experience of participants significantly affects identification of infants language.

We then compared the baseline model with the child interaction model. The result is a chi-square value $X^2(2) = 3.93$, $p = .140$, suggesting the effect of recent childcare and caregiving experience together do not improve model fit significantly (alpha = 0.01). In fact, the AIC value remained at 7633.4 indicating that adding childcare and caregiving experience as predictors caused no improvement to the baseline model. These results indicate that neither recent childcare nor caregiving experience has a significant effect on identifying infant language.

Next, we added gender as another predictive variable into the child interaction model and we called it the gender model (stage two model). Results for this model showed no effect of childcare ($\beta = 0.005$, $z = 1.81$, $p = .071$), caregiving ($\beta = -0.004$, $z = -0.91$, $p = .365$), or gender (β

= -0.048, z = -0.84, p = .404). The AIC of the gender model was 7634.7.

Finally, we compared the gender model with the child interaction model. The result of comparison was not significant, chi-square value $X^2(1) = 0.70$, p = .404. And the AIC value between the two models is increasing. These results indicate that adding gender to our model does not significantly improve the predictive potential of the model. Therefore we can conclude that participant gender does not have a significant effect on identifying infant language.

Discussion

Infant vocalization is an important aspect of the infant language development process and adults' ability to determine the infant's characteristics from their vocalizations may be important for the infant-adult communication domain. Furthermore, variables such as adult listener's sex, and infant caregiving experience might affect the caregiver's ability to communicate with infants. In this study, we measured whether English-speaking monolingual adults can identify the sex, language, and age of an infant after listening to short audio recordings of infant vocalizations and examined the factors that might have an effect on their ability to determine the infant variables.

Contrary to our initial hypothesis, we found that participants were slightly worse than chance at identifying an infant's sex. However, we did find that participants could identify whether the language that the infant was learning was English or non-English better than random chance. Participant's were also able to identify the infant's age range much better than chance. Our exploratory analyses indicated that neither childcare nor caregiving experience significantly affected the ability of the participants to identify the infants' age or language. Nor did participant's gender have an effect on identifying the infants' age or language.

The current study examined adult participants' ability to identify characteristics like age, sex and the language used in the infant's environment using infant vocalizations. Results from our study are consistent with previous findings that adults are able to identify the relative age and primary language of an infant (Olney & Scholnick, 1976; de Boysson-Bardies et al., 1984). Adults' ability to identify infant characteristics from their vocalizations could have an important impact on adult-infant interactions. Previous research shows that infants' vocalizations mature in response to sophisticated adult vocalizations (Goldstein & Schwade, 2008). Our findings, therefore, suggest that adults use identifying vocal cues to facilitate appropriate and contingent responding to enhance language development in infants. As prior research has not analyzed listener's ability to identify infants sex from vocalizations, our study makes a novel contribution to the existing literature on language development. Contrary to our expectations, our findings suggest that participants were not able to identify infants' sex better than chance. We hypothesize that this is the case because there is little to no discernible difference between the sounds from a male or female infant within these age ranges. Also contrary to prior research findings, we found that adult participants with childcare/caregiving experience did not perform better than participants who had no childcare/caregiving experience, at identifying infants' age or language. Our results support Lindová et al. (2015) who found that participants' gender does not affect their ability to accurately identify infant characteristics.

One minor limitation was that, due to the COVID-19 pandemic, we were not able to conduct our research as originally planned, by means of in-person data collection. Therefore, we had to adapt our experiment to be run online. However, we have no reason to believe that this adversely affected our results, as online data collection has been widely validated in a large number of different psychological tasks (Crump et al., 2013). In terms of our exclusion criteria, 166 participants had to be excluded from our sample. However, exclusion of these participants was integral, in particular in relation to our language question, as the knowledge of Spanish, Tsimane', Yêlí-Dnye, Tseltal Mayan, or Quechua could have impacted participants' ability to identify these other languages, causing this subgroup of participants to systematically vary in their responses.

Although we ended up adapting our experiment to be run online, we did not end up posting it outside of the University of Manitoba. Thus, as with a large amount of other research, this poses limits to the generalizability of our research findings. In particular, our sample likely consists of WEIRD subjects (Henrich et al., 2010). WEIRD (Western, educated, and from industrialized, rich, and democratic countries) subjects are typically composed of university-aged samples that have demographic attributes that are fairly narrow in range, in terms of age, socioeconomic status, and cultural norms, where subjects from the University of Manitoba certainly fit this description. Therefore, future directions could include expanding our data collection to platforms such as Amazon Mechanical Turk, or to mirror other large collaborative efforts, such as the ManyLabs and ManyBabies projects (Klein et al, 2014; Bergmann et al., 2017), as a way to include other research labs and diversify our sample.

Another limitation is in regards to the materials we used, particularly in relation to our research design. We are grateful to the BabbleCor team (Meg Cychosz, Amanda Seidl, Elika Bergelson, Marisa Casillas, Gladys Baudet, Anne S. Warlaumont, Camila Scaff, Lisa Yankowitz, and Alejandrina Cristia) for the development of and providing us with the BabbleCor dataset (Cychosz et al., 2020). However, the number of research questions we set out to investigate (e.g.

can participants identify the age, sex, and language of infants) and the available data in the BabbleCor database did not allow us to subset our materials in a way that allowed us to run a full factorial design. We accounted for this by setting a more stringent alpha rate to 0.01 to adjust for the number of questions we were asking, however, future efforts could include contributing to and deriving a larger database of infant vocalizations, consistent with efforts such as the citizen science project hosed on Zooniverse: “Maturity of Baby Sounds” (https://www.zooniverse.org/projects/chiarasemenzin/maturity-of-baby-sounds). This would not only satisfy needs such as our own, but would make materials available to other interested researchers as well.

This research represents a novel contribution to the field of developmental science as little research has been done examining adult judgements of infant vocalizations. This literature is important because it allows us to examine the dynamic of caregiver-infant vocal interactions in a way that has been largely neglected; from the perspective of an adult. The judgements adults make about infant language and age range are important, as language development in infants is a dynamic social process, and, as adults modify their speech in response to infant vocalizations (Elmlinger et al., 2019), development is facilitated. This study contributes to this body of literature by demonstrating that as a whole, adults are accurately able to make these judgements about infant vocalizations, which has never before been demonstrated.

References

Adani, S., & Cepanec, M. (2019). Sex differences in early communication development: Behavioral and neurobiological indicators of more vulnerable communication system development in boys. *Croatian Medical Journal*, *60*(2), 141–149. https://doi.org/10.3325/cmj.2019.60.141

Alhaidary, A., & Rvachew, S. (2018). Cross-linguistic differences in the size of the infant vowel space. *Journal of Phonetics*, *71*, 16–34. https://doi.org/10.1016/j.wocn.2018.07.003

Atkinson, K., & Others (1968). *An experiment on the recognition of babbling* (ED034966). ERIC. https://files.eric.ed.gov/fulltext/ED034966.pdf

Bergmann, C., Frank, M. C., Gonzalez, N., Bergelson, E., Cristia, A., Ferguson, B., Kline, M., Soderstrom, M., Yurovsky, D., Byers-Heinlein, K., Panneton, R., Floccia, C., Lew-Williams, C., Hamlin, K., Shukla, M. (2017). *ManyBabies 1: Infant-directed speech preference*. OSF. Retrieved April 11, 2020, from https://osf.io/re95x/

Best, C. T., Goldstein, L. M., Nam, H., & Tyler, M. D. (2016). Articulating what infants attune to in native speech. *Ecological Psychology*, *28*(4), 216–261. https://doi.org/10.1080/10407413.2016.1230372

Brown, R. (1958). *Words and things*. Free Press.

Cristia, A., Semenzin, C., & Casillas, M. (n.d.) *The maturity of baby sounds.* Zooniverse. Retrieved April 12, 2020, from https://www.zooniverse.org/projects/chiarasemenzin/maturity-of-baby-sounds

Crump, M. J., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PloS one*, *8*(3).

https://doi.org/10.1371/journal.pone.0057410

Cychosz, M., Cristia, A., Bergelson, E., Casillas, M., Baudet, G., Warlaumont, A. S., Scaff, C., Yankowitz, L., & Seidl, A. (2020). *Canonical babble development in a large-scale crosslinguistic corpus*. PsyArXiv. Retrieved March 14, 2020, from https://doi.org/10.31234/osf.io/9vzs5

Cychosz, M., Seidl, A., Bergelson, E., Casillas, M., Baudet, G., Warlaumont, A. S., Scaff, C., Yankowitz, L., & Cristia, A. (2019). *BabbleCor: A crosslinguistic corpus of babble development in five languages* [Data set]. OSF. https://doi.org/10.17605/OSF.IO/RZ4TX

de Boysson-Bardies, B., Halle, P., Sagart, L., & Durand, C. (1989). A crosslinguistic investigation of vowel formants in babbling. *Journal of Child Language*, *16*(1), 1–17. https://doi.org/10.1017/S0305000900013404

de Boysson-Bardies, B., Sagart, L., & Durand, C. (1984). Discernible differences in the babbling of infants according to target language. *Journal of Child Language*, *11*(1), 1–15. https://doi.org/10.1017/S0305000900005559

DePaolis, R. A., Vihman, M. M., & Kunnari, S. (2008). Prosody in production at the onset of word use: A cross-linguistic study. *Journal of Phonetics*, *36*(2), 406–422. https://doi.org/10.1016/j.wocn.2008.01.003

Eimas, P. D., Cooper, W. E., & Corbit, J. D. (1973). Some properties of linguistic feature detectors. *Perception & Psychophysics*, *13*(2), 247–252. https://doi.org/10.3758/BF03214135

Elmlinger, S. L., Schwade, J. A., & Goldstein, M. H. (2019). The ecology of prelinguistic vocal learning: Parents simplify the structure of their speech in response to babbling. *Journal of Child Language, 46*(5), 998–1011. https://doi.org/doi:10.1017/S0305000919000291

Gladding, S. T. (1979). Effects of training versus non-training in identification of infant cry-signals: A longitudinal study. *Perceptual and Motor Skills*, *48*(3), 752–754. https://doi.org/10.2466/pms.1979.48.3.752

Goldstein, M. H., & Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychological Science*, *19*(5), 515–523. https://doi.org/10.1111/j.1467-9280.2008.02117.x

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Hiraoka, D., Ooishi, Y., Mugitani, R., & Nomura, M. (2019). Differential effects of infant vocalizations on approach–avoidance postural movements in mothers. *Frontiers in Psychology*, *10*, Article 1378. https://doi.org/10.3389/fpsyg.2019.01378

Jakobson, R. (1968). *Child language, aphasia, and phonological universals* (A. R. Keiler, Trans.; 3rd ed.). Mouton Publishers. https://doi.org/10.1515/9783111353562 (Original work published in 1941)

Kent, R. D., & Murray, A. D. (1982). Acoustic features of infant vocalic utterances at 3, 6, and 9 months. *The Journal of the Acoustical Society of America*, *72*(2), 353–365. https://doi.org/10.1121/1.388089

Klein, R. A., Ratliff, K. A., Vianello, M., Adams, R. B., Jr., Bahník, Š., Bernstein, M. J., Bocian, K., Brandt, M. J., Brooks, B., Brumbaugh, C. C., Cemalcilar, Z., Chandler, J., Cheong, W., Davis, W. E., Devos, T., Eisner, M., Frankowska, N., Furrow, D., Galliani, E. M., … Nosek, B. A. (2014). Investigating variation in replicability. *Social Psychology*, *45*(3), 142–152. https://doi.org/10.1027/1864-9335/a000178

Lee, C.-C., Jhang, Y., Chen, L., Relyea, G., & Oller, D. K. (2017). Subtlety of ambient-language effects in babbling: A study of English- and Chinese-learning infants at 8, 10, and 12 months. *Language Learning and Development*, *13*(1), 100–126. https://doi.org/10.1080/15475441.2016.1180983

Lindová, J., Špinka, M., & Nováková, L. (2015). Decoding of baby calls: Can adult humans identify the eliciting situation from emotional vocalizations of preverbal infants? *PLOS ONE*, *10*(4), e0124317. https://doi.org/10.1371/journal.pone.0124317

Olney, R. L., & Scholnick, E. K. (1976). Adult judgments of age and linguistic differences in infant vocalization. *Journal of Child Language*, *3*(2), 145–155. https://doi.org/10.1017/S0305000900001410

Parsons, C. E., Young, K. S., Petersen, M. V., Elmholdt, E.-M., J., Vuust, P., Stein, A., & Kringelbach, M. L. (2017). Duration of motherhood has incremental effects on mothers' neural processing of infant vocal cues: A neuroimaging study of women. *Scientific Reports*, *7*, Article 1727 (2017). https://doi.org/10.1038/s41598-017-01776-3

Petitto, L. A., Holowka, S., Sergio, L. E., Levy, B., & Ostry, D. J. (2004). Baby hands that move to the rhythm of language: Hearing babies acquiring sign languages babble silently on the hands. *Cognition*, *93*(1), 43–73. https://doi.org/10.1016/j.cognition.2003.10.007

Quast, A., Hesse, V., Hain, J., Wermke, P., & Wermke, K. (2016). Baby babbling at five months linked to sex hormone levels in early infancy. *Infant Behavior and Development*, *44*, 1–10. https://doi.org/10.1016/j.infbeh.2016.04.002

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

RStudio Team. (2019). *RStudio: Integrated Development Environment for R*. RStudio, Inc. http://www.rstudio.com/.

Snow, D. (2006). Regression and reorganization of intonation between 6 and 23 months. *Child Development*, *77*(2), 281–296. https://doi.org/10.1111/j.1467-8624.2006.00870.x

Stark, R. E., Bernstein, L. E., & Demorest, M. E. (1993). Vocal communication in the first 18 months of life. *Journal of Speech, Language, and Hearing Research*, *36*(3), 548–558. https://doi.org/10.1044/jshr.3603.548

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children’s achievement of language milestones. *Child Development, 72*(3), 748–767. https://doi.org/10.1111/1467-8624.00313

Warlaumont, A. S., Richards, J. A., Gilkerson, J., & Oller, D. K. (2014). A social feedback loop for speech development and its reduction in autism. *Psychological Science*, *25*(7), 1314–1324. https://doi.org/10.1177/0956797614531023