

By 4.5 months, linguistic experience already affects infants' talker processing abilities

Natalie Fecher and Elizabeth K. Johnson*

Department of Psychology, University of Toronto, 3359 Mississauga Road, Mississauga,
Ontario L5L 1C6, Canada

Keywords: infant speech perception, language, indexical processing, talker recognition, language familiarity effect, visual fixation procedure

Running head: Talker recognition in 4.5-month-olds

* **Corresponding author:** elizabeth.johnson@utoronto.ca (E. K. Johnson)

Acknowledgements: Thanks to Lisa Hotson, the members of the Child Language and Speech Studies lab, and all participating families. This research was supported by grants awarded to E.K.J. from SSHRC (Social Sciences and Humanities Research Council of Canada), NSERC (Natural Sciences and Engineering Research Council of Canada), and the CRC (Canada Research Chairs) Program. Portions of this work were presented at the 58th Annual Meeting of the Psychonomic Society in Vancouver, BC, Canada, November 2017.

Abstract

Contemporary models of adult speech perception acknowledge that the processing of linguistic and non-linguistic aspects of the speech signal are interdependent. But when in development does this interdependence first emerge? In the adult literature, one way to demonstrate this relationship has been to examine how language experience affects talker identification. Thus, in the current study, 4- to 5-month-old infants ($N = 96$) were tested on their ability to tell apart talkers in a familiar language (English) compared to unfamiliar languages (Polish or Spanish). Infants readily distinguished between talkers in the familiar language but not in the unfamiliar languages, supporting the hypothesis that the integrated processing of linguistic and non-linguistic information in speech is early-emerging and robust.

In the past, language knowledge and other aspects of perception and cognition were thought to be quite independent (e.g., Fodor, 1983). But we now know that even in early childhood, there is a tight coupling between language development and the development of a wide range of perceptual and cognitive abilities. For example, skill in following eye gaze and general-purpose statistical learning are linked to language abilities (Baldwin, 1993; Brooks & Meltzoff, 2005; Kidd, 2012), and language knowledge supports the recognition of people's voices (Fecher & Johnson, 2018b; Levi & Schwartz, 2013). The current study centres on the last example, i.e., the relationship between language experience and talker recognition, asking when in development this relationship can be observed. Our findings demonstrate a robust and early-emerging influence of linguistic knowledge on talker identity processing.

Research has shown that adults process linguistic aspects of speech (e.g., phonemes, words) and non-linguistic aspects (e.g., talker identity, emotional state) in an integrated manner. For example, language familiarity facilitates vocal emotion detection (Pell & Skorup, 2008; Scherer, Banse, & Wallbott, 2001), consonant and vowel perception is influenced by talker identity (Eisner & McQueen, 2005; Johnson, Strand, & Johnson, 1999), talker familiarity enhances word recognition in adverse listening conditions (Creel, Aslin, & Tanenhaus, 2008; Goldinger, 1996; Mullennix & Pisoni, 1990; Nygaard, Sommers, & Pisoni, 1994), and certain speech sounds carry more information about talker identity than others (Andics, McQueen, & Van Turenout, 2007; Cutler, Andics, & Fang, 2011). And most importantly for the current study, adults recognize talkers better in a familiar language than in an unfamiliar language, a phenomenon known as the language familiarity effect (Goggin, Thompson, Strube, & Simental, 1991; Johnson, Bruggeman, & Cutler, 2018; Levi, 2018b; Orena, Theodore, & Polka, 2015; Perrachione, 2017).

While much evidence suggests that linguistic processing and non-linguistic processing of speech are heavily interactive in adults, our knowledge of when this

bidirectional interaction emerges in childhood is very limited. On the one hand, the two domains might initially develop independently. If this were the case, then we would predict that language-specific knowledge simply refines children's talker recognition skills later in life, with the processing dependencies seen in adults only emerging once children have reached more mature levels of processing in both domains. On the other hand, language and indexical (i.e., talker-related) processing might be contingent on one another from the earliest stages of development. If this were true, then we would predict that children's emerging talker recognition abilities are integrally related to their emerging language abilities. Does the infant and early childhood literature lend support to either of these developmental trajectories?

The talker recognition literature has consistently shown that children's ability to tell apart and recognize talkers continues to improve until late childhood (e.g., Bartholomeus, 1973; Creel & Jiménez, 2012; Mann, Diamond, & Carey, 1979). And of course, language abilities also follow a protracted period of development, with language skills continuously improving until the later childhood years (e.g., Hazan & Barrett, 2000; Nittrouer, 1992; Ohde & Haley, 1997). But although the fact that both talker recognition and language acquisition show protracted development is consistent with the notion that the two abilities influence one another early on, very little work has addressed this idea – and no work has examined this idea in infants under 7.5 months of age. The work that exists in this area suggests that children might perceptually integrate linguistic and indexical information in speech relatively early on. For example, school-age children have difficulty selectively attending to either semantic or talker-specific speech cues (e.g., they more likely report hearing a female voice when listening to a male voice say 'mommy'; Jerger, Martin, & Pirozzolo, 1988; see also Borovsky & Creel, 2014); and infants, too, process talker-related phonetic variation alongside linguistic information during word learning and recognition (Barker & Newman, 2004;

Houston & Jusczyk, 2000; Rost & McMurray, 2009). Moreover, experience with a greater variety of speech sounds has been proposed as an explanation for why infants growing up in bilingual households are better at face-voice matching in an unfamiliar language than are monolingual infants (Fecher & Johnson, 2018a; see Levi, 2018a, for relevant work with older children).

Additional evidence for the early integration of language knowledge and talker recognition comes from studies showing that children display a language familiarity effect long before their language abilities are fully developed. For example, 5- to 15-year-old children show a native-language benefit in a talker processing task (Fecher & Johnson, 2018b; Levi, 2018a; Levi & Schwartz, 2013; Perea et al., 2014), and even infants exhibit this effect in a talker discrimination task (Fecher & Johnson, 2018c; Johnson, Westrek, Nazzi, & Cutler, 2011). The two infant studies published in this area found that 7.5-month-old infants could tell apart the voices of female talkers when the talkers spoke a familiar language; however, infants no longer succeeded at this task when the same talkers spoke an unfamiliar language. But do these findings necessarily imply that language and talker processing interact from the earliest stages of development?

The finding that language knowledge affects talker discrimination as early as 7.5 months of age supports the idea that linguistic and non-linguistic processing of speech interact quite early on. However, 7.5-month-olds have already gathered much linguistic knowledge about their native language and they are by no means linguistically naïve. By 7–8 months, infants comprehend a surprisingly large number of words (e.g., Bergelson & Swingley, 2012; Tincoff & Jusczyk, 1999), and they demonstrate considerable attunement to the phoneme inventory and prosodic structure of their native language (e.g., Johnson, 2016; Maurer & Werker, 2014). It is thus conceivable that the language familiarity effect only emerges once infants have reached 7 or 8 months of age and have started to build a

phonological inventory and receptive lexicon in their native language. The currently-available data are hence consistent with two divergent possibilities: either the integration of linguistic and non-linguistic information in speech starts at 7 to 8 months of age, or the two types of information are integrated even earlier in development – perhaps even from birth. To distinguish between these possibilities, we must test younger infants who are less perceptually attuned to their native language.

In the current study, we tested infants between 4 and 5 months of age on talker discrimination in their native language compared to two different foreign languages. If linguistic and indexical processing interact very early in development, then 4.5-month-olds should be better at telling apart native-language than foreign-language talkers. However, if language and talker processing interact later in development, i.e., at an age where infants are more substantially tuned in to their native language, then 4.5-month-olds should perform similarly across languages.

Method

Participants. Ninety-six full-term monolingual English-learning 4- to 5-month-old infants ($M_{age} = 136$ days, range = 120–152; 48 female) from the Greater Toronto Area were tested. Infants were exposed to English at least 90% of the time, and none of them had prior experience with Polish or Spanish (the two foreign languages included in this study). The data for 43 additional infants were excluded from data analysis due to failure to complete at least six exposure trials before reaching a predefined exposure criterion and proceeding to test (11 infants in the native-language condition, 9 infants in the foreign-language condition), fussing (5 for native, 5 for foreign), failure to reach posttest criterion (4 for native, 4 for foreign), or if infants completed the maximum number of exposure trials (4 for native, 1 for foreign). Exclusion criteria and sample size were based on past infant research on talker

processing (Fecher & Johnson, 2018a, c; Johnson et al., 2011). Infants were tested on two sets of voices and two foreign languages to ensure replicability and generalizability of our findings (see also Design section).

Stimuli. The materials consisted of 40 sentences in each of English and Polish (see Fecher & Johnson, 2018c) as well as Spanish (see Paquette-Smith & Johnson, 2015). Examples of these sentences are shown in the Appendix. The sentences were read in an adult-directed manner by four female bilingual speakers ($M_{\text{age}} = 21.0$ years, $SD = 2.2$). Two of the speakers were English-Polish bilinguals; they recorded the English and Polish sentences (these were the same recordings as those used in Fecher & Johnson, 2018c). The other two speakers were English-Spanish bilinguals; they recorded the English and Spanish sentences (the Spanish recordings were the same as used in Fecher & Johnson, 2018a). All speakers learned both their first and second language from birth, all regularly speak both languages, and none of them had a detectable foreign accent when speaking English. The two sets of English recordings were used as native-language stimuli, and the Polish and Spanish recordings served as foreign-language stimuli. Thirty-six sentences per set were used during the exposure phase of the experiment, and the remaining four sentences were used at test.

Procedure. Infants were tested using the visual fixation procedure. The experimental setup and procedure were identical to those used in Fecher & Johnson (2018c). Infants sat on their caregiver's lap in a double-walled, sound attenuated Industrial Acoustics Company booth facing a 21.5-inch computer monitor that showed a multi-colored flickering checkerboard. Stimuli were played through Alesis M1 Active 520 loudspeakers at a constant, comfortable listening level. The experimenter relayed infants' looking behavior to a computer outside the booth. Trials were initiated by the experimenter once the infant had oriented toward a blinking red star that was presented before each trial in the center of the monitor. Caregivers listened to masking music mixed with speech stimuli from the experiment and

played through noise-canceling headphones to prevent them from influencing their child's performance.

The experiment consisted of an exposure phase, where infants were familiarized with one of the talkers, and a test phase, where infants were tested on their ability to notice a talker change. In each infant-controlled exposure trial (maximum 16 s long), infants heard repetitions of either two native-language or two foreign-language sentences (depending on language condition) spoken by the same talker. The exposure phase ended once infants' looking time had decreased to 65% of the initial looking time (calculated over a sliding window of three trials) or once they had completed the maximum of 18 exposure trials. At test, infants were presented with two 'same-voice' trials (where they heard the familiar voice one more time) and two 'different-voice' trials (where they heard an unfamiliar voice). In a pretest trial (before the first exposure trial) and a posttest trial (after the last test trial), infants saw a colorful spinning windmill (combined with a sound effect), which acted as a control for whether infants had become uninterested in the stimuli or generally fatigued. If the looking time during posttest reached at least 80% of the looking time during pretest, then we would infer that infants were still attentive to the task and therefore the data were usable. In keeping with Fecher & Johnson (2018c), we would exclude data if infants did not reach the 80% posttest criterion, if infants completed all 18 exposure trials, or if infants did not complete at least six exposure trials before proceeding to test. We predicted that if infants can tell the two voices apart, then their mean looking time during same-voice trials and their mean looking time during different-voice trials should significantly differ.

Design. Infants were randomly assigned to one of two language conditions: native language (English) or foreign language (Polish or Spanish). In the native-language condition ($N = 48$), half of the infants ($N = 24$) were tested on talker pair 1 (i.e., the two English-Polish bilinguals speaking English) and the other half ($N = 24$) were tested on talker pair 2 (i.e., the

two English-Spanish bilinguals speaking English). In the foreign-language condition ($N = 48$), half of the infants ($N = 24$) were tested on Polish (i.e., the two English-Polish bilinguals speaking Polish) and the other half ($N = 24$) were tested on Spanish (i.e., the two English-Spanish bilinguals speaking Spanish). Across infants, we counterbalanced the order of presentation of the two types of test trials, which talker was heard during exposure and which one was heard during test, and which sentences were heard during exposure and test. We predicted that infants should show a looking time difference between same-voice and different-voice trials (and thus evidence of talker discrimination) if assigned to the English condition, but that infants' performance for either foreign language (i.e., Polish or Spanish) would critically depend on whether or not linguistic and non-linguistic aspects of speech are processed in an integrated manner at 4–5 months. That is, if the processing of language and talker information are interdependent early on (i.e., before infants have developed a substantial proto-lexicon), then we should see improved talker discrimination in English than in Polish or Spanish. If, however, this interdependence develops later in infancy (i.e., at around 7.5 months), then infants should perform similarly across language conditions.

Results

Infants assigned to the native-language condition and those assigned to the foreign-language condition did not differ in the number of exposure trials completed before proceeding to test ($M_{native} = 9.3$, $SD = 2.6$; $M_{foreign} = 9.5$, $SD = 3.1$), $t(94) = -.32$, $p = .749$. Moreover, infants did not differ between language conditions in their mean looking times (in seconds) during the first three exposure trials ($M_{native} = 14.7$, $SD = 2.2$; $M_{foreign} = 14.4$, $SD = 2.4$), $t(94) = .80$, $p = .428$; and infants also did not differ by condition in their mean looking times during the last three exposure trials ($M_{native} = 8.0$, $SD = 1.7$; $M_{foreign} = 8.0$, $SD = 1.7$), $t(94) = .17$, $p = .862$. This indicates that infants' looking times during exposure to the first

talker (and thus their attention paid to the first talker) did not differ across language conditions and this factor therefore cannot account for any performance differences seen at test.

To assess infants' talker discrimination abilities, we compared mean looking time (in seconds) during same-voice trials to mean looking time during different-voice trials. A linear mixed-effects regression (LMER) model was conducted using the lme4 package for R (Bates, Mächler, Bolker, & Walker, 2015), with mean looking time as the dependent variable and contrast-coded fixed effects for voice (same, different), language (native, foreign), talker pair (pair 1, pair 2), and their interactions. The model also included a random intercept for participant and a random slope for voice by participant. Model comparisons were performed to determine whether the inclusion of each fixed factor and the interactions made a significant contribution to the model. No significant main effects of voice, language, or talker pair were obtained, and with one exception, none of the interactions were significant ($\chi^2 < 2.54, p > .111$). The absence of an effect of talker pair (including its interactions) suggests that infants performed similarly across the two pairs of talkers and foreign languages. Critically, however, the voice \times language interaction was significant ($\beta = -2.23, SE = 0.90, \chi^2(1, N = 96) = 5.88, p = .015$). This indicates that the difference in looking time between same-voice and different-voice trials was dependent on whether infants had been tested on their native language or on a foreign language (regardless of the type of foreign language).

To further assess the effect of language on infants' looking times during same-voice and different-voice trials, a separate model was constructed per language condition, with each model including a fixed effect for voice, a random intercept for participant, and a by-participant random slope for voice. In the native-language condition, infants looked significantly longer during same-voice ($M = 8.0, SD = 4.4$) than different-voice ($M = 6.2, SD = 3.1$) test trials ($\beta = -1.84, SE = 0.54, \chi^2(1, N = 48) = 10.91, p < .001$, suggesting that

infants successfully discriminated between the talkers in English (see Figure 1, two leftmost bars; for discussion of the direction of preference see the Discussion). In the foreign-language condition, however, mean looking times during same-voice ($M = 7.7$, $SD = 4.5$) and different-voice ($M = 8.0$, $SD = 4.2$) test trials did not significantly differ ($p = .603$), which indicates that infants could not tell apart the talkers in Polish or Spanish (see Figure 1, two rightmost bars).

Discussion

Past work has shown that language experience affects talker processing in the latter half of the first year of life, i.e., at an age where infants first show comprehension of commonly heard words and where they have made substantial progress tuning in to the segmental and prosodic structure of their native language. In the current study, we show that language knowledge influences talker processing far earlier than any other study has shown. We found that English-learning 4.5-month-olds were better at talker discrimination in a familiar language (English) than in unfamiliar languages (Polish or Spanish). This finding strongly supports the hypothesis that language learning and the development of non-linguistic cognitive abilities (such as talker processing) are fundamentally linked from an early age.

In addition to being the first study to demonstrate a link between language and talker discrimination in infants under 5 months of age, our findings contribute more generally to the language learning and talker recognition literature. In the past, the infant literature on talker recognition has nearly exclusively focused on infants' perception of familiar talkers (e.g., DeCasper & Fifer, 1980; Mehler, Bertoncini, Barrière, & Jassik-Gerschenfeld, 1978; Mills & Melhuish, 1974) and acoustically-distinct (male versus female) talkers (e.g., Floccia, Nazzi, & Bertoncini, 2000; Lecanuet, Granier-Deferre, Jacquet, Capponi, & Ledru, 1993; Miller, 1983; see Fecher, Paquette-Smith, & Johnson, 2019, for discussion). Here, we present first

evidence that by 5 months, infants can detect fine-grained phonetic differences even in the voices of unfamiliar, same-age and same-gender talkers (at least in their native language). Moreover, our finding that 4.5-month-olds are better at talker discrimination in a familiar (than an unfamiliar) language suggests that infants already show similar processing dependencies across linguistic and non-linguistic domains as do older children and adults. From this we conclude that language knowledge does not simply refine talker recognition later in life, i.e., once both domains are fully developed. Instead, talker recognition hinges on linguistic processing very early on, suggesting that both domains are intrinsically related long before infants reach certain linguistic milestones (such as comprehension of first words or mastery of language-specific phonology). Future work should more closely examine the developmental trajectory of this interdependence (e.g., by testing even younger infants). Moreover, studies should address whether the present findings generalize to different testing paradigms (like recognition rather than discrimination; see Fecher et al., 2019), to different talkers (like talkers varying in voice quality, accent, or speaking style), and to different infant populations. Regarding the latter, one might ask whether infants with more or less experience with talker variation (see Bergmann & Cristia, 2018), or infants exposed to one language (or language variety) from birth versus infants raised in multilingual (or multi-accented) homes (see Fecher & Johnson, 2018a), process talker identity in speech differently.

Aside from clear developmental implications for talker recognition, does the early emergence of improved native-language talker processing offer young language learners benefits beyond those associated with the development of mature talker recognition abilities? Attending to voice characteristics that convey talker identity not only helps infants to separate out one talker from another (a prerequisite for successful talker recognition), it may also support language learning more generally. It has been known for some time that sensitivity to talker-related speech cues facilitates the tracking of lower- and higher-level linguistic

structure in the speech stream. For example, infants track phonetic characteristics of different talkers and use this information to infer the referents of novel word labels produced by these talkers (Weatherhead & White, 2016), and infants maximize word recognition efficiency by attuning to the typical speech patterns produced by speakers of different language variants (van der Feest & Johnson, 2016). Interestingly, a recent study on artificial language learning in 12-month-olds found that infants use voice characteristics as a cue to learn talker-dependent grammatical structure (Gonzales, Gerken, & Gómez, 2018). Evidently, processing contextual information about who is talking plays an important role in first language acquisition. We can therefore conceptualize our finding that infants are more proficient at detecting talker differences in their native language (than in foreign languages) as an adaptation to their spoken language environment, whereby reliable (native-language) talker processing will ultimately help to advance their (native) language abilities. And we may even go one step further and speculate that if infants adapt their talker processing strategies to their environment, then the nature and time course of this adaptation process might differ in infants with different linguistic experiences. Indeed, initial work in this area supports this hypothesis, showing that bilingual but not monolingual 9-month-olds succeed at foreign-language talker recognition (Fecher & Johnson, 2018a).

Finally, from a methodological standpoint, it is interesting to note that although the 7.5-month-olds tested previously (Fecher & Johnson, 2018c; Johnson et al., 2011) looked longer during different-voice than same-voice test trials, the 4.5-month-olds tested here looked longer during same-voice than different-voice trials. Why might this be the case? Infant looking preferences are known to be susceptible to the experimenter's choice of exposure criterion, infants' age (along with factors like cognitive maturity), and stimulus complexity (e.g., Houston-Price & Nakai, 2004; Hunter & Ames, 1988; Mather, 2013; Oakes, 2010). In the current study, instead of using the exposure criterion typically used with infants

under 5 months of age (i.e., 50%), we matched the criterion to earlier work on talker discrimination in 7.5-month-olds (i.e., 65%). But even though we used the same criterion, infants at 4.5 and those at 7.5 months responded differently in the exact same task. Thus, the behavioral differences seen across our 4.5- and 7.5-month-old infant populations may be attributed to the age difference and to the fact that the experimental stimuli (human voices) were quite complex. As pointed out by an anonymous reviewer, it is furthermore tempting to speculate that the current finding of a looking time difference between familiar and novel voices (in English) may have been driven by an active avoidance of novel voices rather than a preference for familiar voices (or indeed a combination of both types of responses). Although this is an interesting suggestion, the study was not designed to distinguish between these possibilities – and indeed, infant data are generally interpreted as a preference rather than an avoidance when in fact the data cannot typically distinguish between the two. But setting aside these methodological considerations, the significant performance differences between language conditions in any case clearly show that as early as 4 months of age, infants are better at telling apart talkers in a familiar language than in an unfamiliar language.

In conclusion, we have shown that language and indexical processing of speech are tightly interwoven from the early stages of child development (rather than initially develop independently). We have demonstrated this early dependence between linguistic and non-linguistic processing by showing that by 4–5 months of age, infants are already better at talker discrimination in a familiar language than in an unfamiliar language. This study thus lays a foundation for future research on the relationship between linguistic and indexical processing in infancy and throughout the life span. To move forward in our understanding of the cognitive processes involved in processing the information-rich speech signal produced by the human voice, we may need to rethink how to best capture the relationship between different types of information carried by the speech signal. That is, if talker recognition draws

upon linguistic experience (and vice versa) from day one of human development, then perhaps the separation of linguistic and non-linguistic aspects of the speech signal is somewhat of a false dichotomy. Future work in this area will contribute to our understanding of talker recognition abilities across development, and it will inform theories linking language learning with the development of other cognitive and perceptual domains.

DRAFT ONLY

Appendix: Examples of test sentences*English*

1. The last concert given at the opera was a tremendous success.
2. My grandparents' neighbor is the most charming person I know.

Polish

1. Pociąg odjechał ze stacji więcej niż piętnaście minut temu. (The train left the station more than 15 minutes ago.)
2. Dzisiaj był piękny dzień, więc poszłam na spacer z moją mamą. (Today was a beautiful day, so I went for a walk with my mom.)

Spanish

1. El niño se levantó temprano para ver el sol. (The boy got up early to see the sun.)
2. El nuevo presidente será elegido en mayo. (The new president will be elected in May.)

References

- Andics, A., McQueen, J. M., & Van Turennout, M. (2007). Phonetic context influences voice discriminability. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1829–1832). Saarbrücken, Germany.
- Baldwin, D. A. (1993). Early Referential Understanding: Infants' Ability to Recognize Referential Acts for What They Are. *Developmental Psychology*, *29*, 832–843.
<https://doi.org/10.1037/0012-1649.29.5.832>
- Barker, B. A., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, *94*, 45–53.
<https://doi.org/10.1016/j.cognition.2004.06.001>
- Bartholomeus, B. (1973). Voice identification by nursery school children. *Canadian Journal of Psychology*, *27*, 464–472.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.
<https://doi.org/10.18637/jss.v067.i01>
- Bergelson, E., & Swingle, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 3253–3258. <https://doi.org/10.1073/pnas.1113380109>
- Bergmann, C., & Cristia, A. (2018). Environmental influences on infants' native vowel discrimination: The case of talker number in daily life. *Infancy*, *23*, 484–501.
<https://doi.org/10.1111/infa.12232>
- Borovsky, A., & Creel, S. C. (2014). Children and adults integrate talker and verb information in online processing. *Developmental Psychology*, *50*, 1600–1613.
<https://doi.org/http://dx.doi.org/10.1037/a0035591>
- Brooks, R., & Meltzoff, A. (2005). The development of gaze following and its relations to

- language. *Developmental Science*, 8, 535–543. <https://doi.org/10.1111/j.1467-7687.2005.00445.x>.The
- Creel, S. C., Aslin, R. N., & Tanenhaus, M. K. (2008). Heeding the voice of experience: The role of talker variation in lexical access. *Cognition*, 106, 633–664. <https://doi.org/10.1016/j.cognition.2007.03.013>
- Creel, S. C., & Jiménez, S. R. (2012). Differences in talker recognition by preschoolers and adults. *Journal of Experimental Child Psychology*, 113, 487–509. <https://doi.org/10.1016/j.jecp.2012.07.007>
- Cutler, A., Andics, A., & Fang, Z. (2011). Inter-dependent categorization of voices and segments. In W.-S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 552–555). Hong Kong, China.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208, 1174–1176.
- Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, 67, 224–238.
- Fecher, N., & Johnson, E. K. (2018a). Bilingual infants excel at foreign-language talker recognition. *Developmental Science*, e12778. <https://doi.org/https://doi.org/10.1111/desc.12778>
- Fecher, N., & Johnson, E. K. (2018b). Effects of language experience and task demands on talker recognition by children and adults. *The Journal of the Acoustical Society of America*, 143, 2409–2418. <https://doi.org/10.1121/1.5032199>
- Fecher, N., & Johnson, E. K. (2018c). The native-language benefit for talker identification is robust in 7.5-month-old infants. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44, 1911–1920. <https://doi.org/10.1037/xlm0000555>
- Fecher, N., Paquette-Smith, M., & Johnson, E. K. (2019). Resolving the (apparent) talker

recognition paradox in developmental speech perception. *Infancy*.

<https://onlinelibrary.wiley.com/doi/full/10.1111/infa.12290>

Floccia, C., Nazzi, T., & Bertoncini, J. (2000). Unfamiliar voice discrimination for short stimuli in newborns. *Developmental Science*, *3*, 333–343. <https://doi.org/10.1111/1467-7687.00128>

Fodor, J. A. (1983). *Modularity of mind*. Cambridge, MA: MIT Press.

Goggin, J. P., Thompson, C. P., Strube, G., & Simental, L. R. (1991). The role of language familiarity in voice identification. *Memory & Cognition*, *19*, 448–458. <https://doi.org/10.3758/BF03199567>

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1166–1183. <https://doi.org/10.1037/0278-7393.22.5.1166>

Gonzales, K., Gerken, L., & Gómez, R. L. (2018). How who is talking matters as much as what they say to infant language learners. *Cognitive Psychology*, *106*, 1–20. <https://doi.org/10.1016/j.cogpsych.2018.04.003>

Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*, *28*, 377–396. <https://doi.org/10.1006/jpho.2000.0121>

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*, 341–348. <https://doi.org/10.1002/icd.364>

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 1570–1582. <https://doi.org/10.1037/0096-1523.26.5.1570>

Hunter, M. A., & Ames, E. W. (1988). A multifactor model of infant preferences for novel and familiar stimuli. In C. Rovee-Collier & L. P. Lipsitt (Eds.), *Advances in Infancy*

- Research* (Vol. 5, pp. 69–95). Stamford, CT: Ablex.
- Jerger, S., Martin, R. C., & Pirozzolo, F. J. (1988). A developmental study of the auditory Stroop effect. *Brain and Language*, *35*, 86–104.
- Johnson, E. K. (2016). Constructing a proto-lexicon: An integrative view of infant language development. *Annual Review of Linguistics*, *2*, 391–412.
<https://doi.org/10.1146/annurev-linguistics-011415-040616>
- Johnson, E. K., Bruggeman, L., & Cutler, A. (2018). Abstraction and the (misnamed) language familiarity effect. *Cognitive Science*, *42*, 633–645.
<https://doi.org/10.1111/cogs.12520>
- Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science*, *14*, 1002–1011.
<https://doi.org/10.1111/j.1467-7687.2011.01052.x>
- Johnson, K., Strand, E. A., & Johnson, K. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, *27*, 359–384.
<https://doi.org/10.1006/jpho.1999.0100>
- Kidd, E. (2012). Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology*, *48*, 171–184. <https://doi.org/10.1037/a0025405>
- Lecanuet, J.-P., Granier-Deferre, C., Jacquet, A.-Y., Capponi, I., & Ledru, L. (1993). Prenatal discrimination of a male and a female voice uttering the same sentence. *Early Development and Parenting*, *2*, 217–228.
- Levi, S. V. (2018a). Another bilingual advantage? Perception of talker-voice information. *Bilingualism: Language and Cognition*, *21*, 523–536.
<https://doi.org/10.1017/S1366728917000153>
- Levi, S. V., & Schwartz, R. G. (2013). The development of language-specific and language-independent talker processing. *Journal of Speech, Language, and Hearing Research*, *56*,

- 913–920. [https://doi.org/10.1044/1092-4388\(2012/12-0095\)](https://doi.org/10.1044/1092-4388(2012/12-0095))
- Levi, S. V. (2018b). Methodological considerations for interpreting the Language Familiarity Effect in talker processing. *Wiley Interdisciplinary Reviews: Cognitive Science*, 10, 1–15. <https://doi.org/10.1002/wcs.1483>
- Mann, V. A., Diamond, R., & Carey, S. (1979). Development of voice recognition: Parallels with face recognition. *Journal of Experimental Child Psychology*, 27, 153–165. [https://doi.org/10.1016/0022-0965\(79\)90067-5](https://doi.org/10.1016/0022-0965(79)90067-5)
- Mather, E. (2013). Novelty, attention, and challenges for developmental psychology. *Frontiers in Psychology*, 4, 491. <https://doi.org/10.3389/fpsyg.2013.00491>
- Maurer, D., & Werker, J. F. (2014). Perceptual narrowing during infancy: A comparison of language and faces. *Developmental Psychobiology*, 56, 154–178. <https://doi.org/10.1002/dev.21177>
- Mehler, J., Bertoncini, J., Barrière, M., & Jassik-Gerschenfeld, D. (1978). Infant recognition of mother's voice. *Perception*, 7, 491–497. <https://doi.org/10.1068/p070491>
- Miller, C. L. (1983). Developmental changes in male/female voice classification by infants. *Infant Behavior and Development*, 6, 313–330. [https://doi.org/10.1016/S0163-6383\(83\)80040-X](https://doi.org/10.1016/S0163-6383(83)80040-X)
- Mills, M., & Melhuish, E. (1974). Recognition of mother's voice in early infancy. *Nature*, 252, 123–124. <https://doi.org/10.1038/252123a0>
- Mullennix, J. W., & Pisoni, D. B. (1990). Stimulus variability and processing dependencies in speech perception. *Perception & Psychophysics*, 47, 379–390. <https://doi.org/10.3758/BF03210878>
- Nittrouer, S. (1992). Age-related differences in perceptual effects of formant-transitions within syllables and across syllable boundaries. *Journal of Phonetics*, 20, 351–382.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-

contingent process, 5, 42–46. <https://doi.org/https://doi.org/10.1111/j.1467-9280.1994.tb00612.x>

Oakes, L. M. (2010). Using habituation of looking time to assess mental processes in infancy. *Journal of Cognition and Development, 11*, 255–268.

<https://doi.org/10.1080/15248371003699977>

Ohde, R. N., & Haley, K. L. (1997). Stop-consonant and vowel perception in 3- and 4-year-old children. *The Journal of the Acoustical Society of America, 102*, 3711–3722.

<https://doi.org/10.1121/1.420135>

Orena, A. J., Theodore, R. M., & Polka, L. (2015). Language exposure facilitates talker learning prior to language comprehension, even in adults. *Cognition, 143*, 36–40.

<https://doi.org/10.1016/j.cognition.2015.06.002>

Paquette-Smith, M., & Johnson, E. K. (2015). Spanish-accented English is Spanish to English-learning 5-month-olds. In The Scottish Consortium for ICPHS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, United Kingdom.

Pell, M. D., & Skorup, V. (2008). Implicit processing of emotional prosody in a foreign versus native language. *Speech Communication, 50*, 519–530.

<https://doi.org/https://doi.org/10.1016/j.specom.2008.03.006>

Perea, M., Jiménez, M., Suárez-Coalla, P., Fernández, N., Viña, C., & Cuetos, F. (2014). Ability for voice recognition is a marker for dyslexia in children. *Experimental Psychology, 61*, 480–487.

<https://doi.org/10.1027/1618-3169/a000265>

Perrachione, T. K. (2017). Speaker recognition across languages. In S. Frühholz & P. Belin (Ed.), *The Oxford Handbook of Voice Perception*. Oxford: Oxford University Press.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science, 12*, 339–49.

<https://doi.org/10.1111/j.1467-7687.2008.00786.x>

- Scherer, K. R., Banse, R., & Wallbott, H. G. (2001). Emotion inferences from vocal expression correlate across languages and cultures. *Journal of Cross-Cultural Psychology*, *32*, 76–92. <https://doi.org/https://doi.org/10.1177/0022022101032001009>
- Tincoff, R., & Jusczyk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*, 172–175. <https://doi.org/10.1111/1467-9280.00127>
- van der Feest, S. V. H., & Johnson, E. K. (2016). Input-driven differences in toddlers' perception of a disappearing phonological contrast. *Language Acquisition*, *23*, 89–111. <https://doi.org/10.1080/10489223.2015.1047096>
- Weatherhead, D., & White, K. S. (2016). He says potato, she says potahto: Young infants track talker-specific accents. *Language Learning and Development*, *12*, 92–103. <https://doi.org/10.1080/15475441.2015.1024835>

Figure

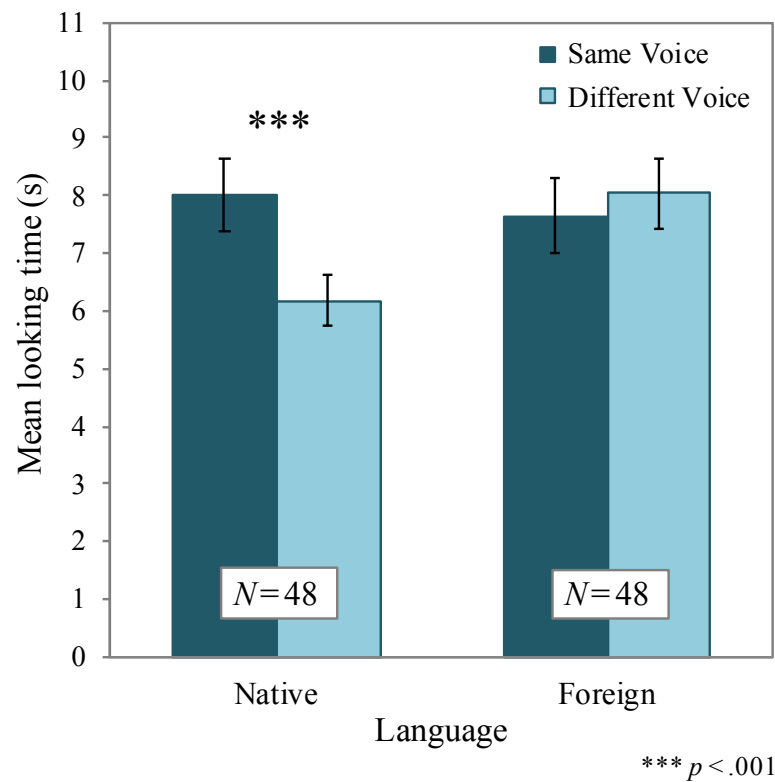


Figure 1. Mean looking times (in seconds) during same-voice and different-voice test trials in the native-language and foreign-language conditions. Infants' looking times during same-voice and different-voice trials significantly differed in the native-language condition but not in the foreign-language condition, suggesting that language experience affects talker processing from a very early stage in child development (4.5 months). Error bars represent standard error of the mean.