The native-language benefit for talker identification is robust in 7.5-month-old infants

Natalie Fecher* and Elizabeth K. Johnson
University of Toronto, Canada

© 2018, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors permission. The final article will be available, upon publication, via its DOI: 10.1037/xlm0000555

* Corresponding author at: Department of Psychology, University of Toronto, 3359 Mississauga Road, Mississauga, ON L5L 1C6, Canada.

Email addresses: natalie.fecher@utoronto.ca, elizabeth.johnson@utoronto.ca
Abstract

Adults recognize talkers better when the talkers speak a familiar language than when they speak an unfamiliar language. This Language Familiarity Effect (LFE) demonstrates the inseparable nature of linguistic and indexical information in adult spoken language processing. Relatively little is known about children’s integration of linguistic and indexical information in speech. For example, to date, only one study has explored the LFE in infants (Johnson, Westrek, Nazzi, & Cutler, 2011). Here, we aim to better understand the maturation of speech processing abilities in infants by replicating Johnson et al. using a more stringent experimental design (eliminating a potential voice/language confound), a different test population (English- rather than Dutch-learning infants), and a new language pairing (English and Polish rather than Dutch and Italian/Japanese). Furthermore, we explore the language exposure conditions required for infants to develop an LFE for a formerly unfamiliar language. Based on previous studies (including the perceptual narrowing literature), we hypothesized that infants might develop an LFE more readily than adults. Although our findings replicate Johnson et al. – demonstrating that the LFE is robust in 7.5-month-olds – we find no evidence that infants need less language exposure than adults to develop an LFE. We conclude that both infants and adults need extensive (potentially live) exposure to an unfamiliar language before talker identification in that language improves. Moreover, the LFE is likely rooted in early-emerging phonology rather than shared lexical knowledge, and infants already closely resemble adults in their processing of linguistic and indexical information.

Keywords: talker recognition, Language Familiarity Effect, infant speech perception, Visual Fixation Procedure, phonological development
Introduction

Adults recognize talkers who speak a familiar language more successfully than talkers who speak an unfamiliar language. This Language Familiarity Effect (LFE) is widely recognized as a cornerstone of adult talker recognition, and shows that the processing of linguistic information (about the speech content) and indexical information (about the talker’s identity) are tightly interwoven in mature speech perception (Remez, Fellowes, & Rubin, 1997). In adults, the effect has been demonstrated across different experimental paradigms and language pairings (e.g., English/Dutch: Johnson, Bruggeman, & Cutler, 2017; English/French: Kadam, Orena, Theodore, & Polka, 2016; English/Korean: Bregman & Creel, 2014; English/Mandarin Chinese: Perrachione, Del Tufo, & Gabrieli, 2011). Studying the developmental roots of the LFE can help us to better understand when, how or why the effect would emerge in adults, and thus enhance our understanding of mature talker recognition abilities and the nature of human speech processing more generally. Here, we examined the emergence of the LFE in infants, and explored what kind of language experience is necessary for infants to show improved talker recognition in an unfamiliar language.

Although the native-language benefit for talker identification is firmly established in the adult literature, developmental work on the LFE is still sparse. To date, only three published studies have examined the role of language familiarity in talker recognition by school-aged children (Levi, 2017; Levi & Schwartz, 2013; Perea et al., 2014). These studies have found that typically-developing children between 6 and 15 years of age recognize talkers better in their native language than a foreign language, but have not established a clear timeline for when and how the LFE develops. In addition, only one published study has tested the LFE in infants, and this study reported that Dutch-learning 7.5-month-old infants were better at telling apart Dutch speakers
than speakers of either Italian or Japanese (Johnson, Westrek, Nazi, & Cutler, 2011). Given infants’ limited lexical processing skills at that age, this finding strongly counters suggestions that the LFE for talker identification is caused by listener’s comprehension of a talker (e.g., Goggin, Thompson, Strube, & Simental, 1991; Perrachione, Dougherty, McLaughlin, & Lember, 2015). Rather, based on this infant finding, it seems far more likely that language-specific phonological knowledge is the main reason that talkers are easier to identify when they speak a familiar language.

But can we be certain that infants really demonstrate the LFE in the latter half of the first year of life? By this age, infants have some knowledge of the prosodic structure of their native language and have started tuning in to some aspects of the segmental structure of their native language (see Johnson, 2016, and Jusczyk, 1997, for review). But at the same time, 7.5-month-olds have not yet fully mastered the phonology of their native language. For example, they do not appear to fully understand position-specific allophones (Jusczyk, Hohne, & Bauman, 1999), are not yet sensitive to the phonotactic structure of their native language (Jusczyk & Luce, 1994), may possess under-specified phonological representations (Bergelson & Swingley, 2017), are overly-reliant on certain word segmentation heuristics (Jusczyk, Houston, & Newsome, 1999), and have difficulties recognizing words across talker and accent variability (Houston & Jusczyk, 2000; see, however, Van Heugten & Johnson, 2012). Given all these limitations in their knowledge of native-language phonology, do 7.5-month-olds really possess enough language-specific knowledge to exhibit an LFE for talker recognition?

In addition to asking whether young infants have sufficient language-specific knowledge to demonstrate the LFE, there are other independent grounds to question the claim that 7.5-month-olds already show a native-language benefit for talker identification. For example, infants’ performance in the talker discrimination task
reported in Johnson et al. (2011) was surprisingly high, and indeed much higher than we would expect based on previous research. While infants and even newborns have been shown to recognize voices under certain favourable conditions (i.e., when the voices are highly familiar, e.g., DeCasper & Fifer, 1980, or when the talkers differ in gender, e.g., Floccia, Nazzi, & Bertoncini, 2000), we also know that even 3- to 6-year-old children still struggle to learn to identify unfamiliar talkers that match in gender (Creel & Jiménez, 2012). If school-aged children have difficulties recognizing unfamiliar same-gender talkers, then how do we account for the finding that the 7.5-month-olds tested in Johnson et al. (2011) performed so well with unfamiliar same-gender speakers?

For all the reasons discussed above, and the fact that Johnson et al. (2011) is the only published report of an LFE in infants, our first goal of the current study was to replicate the original Johnson et al. study using a more stringent experimental design, a different test population and new stimuli. To allow for comparability of the earlier and the current study, we used the same age group and a similar test methodology as Johnson et al. (2011). This age group was originally chosen for studying the LFE in infants because 7.5 months is a reasonable age to expect an LFE to emerge if the LFE does not depend on comprehension (as noted earlier, 7.5-month-olds have an emerging knowledge of native-language phonology, but their comprehension abilities are still very limited). However, in contrast to Johnson et al., our stimuli were produced by bilingual instead of monolingual talkers, which decreased the likelihood that the speakers of one language were more distinct than the speakers of another language (Johnson et al. used different speakers for each language condition, leading to the possibility that the Dutch speakers were simply more distinct – and thus easier to tell apart – than the Italian and Japanese speakers). Note that attempting this replication not
only contributes to our understanding of how children develop adult-level talker recognition and spoken language processing skills, it also addresses recently-expressed concerns regarding replicability in infant research more generally (Frank et al., 2017).

Aside from establishing the validity of the original study reporting an infant LFE, a second goal of the current study was to explore how much and what kind of language exposure infants require to elicit an LFE for a formerly unfamiliar language. For adults, there is an emerging understanding that although advanced language competency and speaking proficiency is not needed to facilitate talker recognition, linguistic experience with a language has to be fairly substantial for talker identification in that language to improve (Bregman & Creel, 2014; Köster & Schiller, 1997). One study found that English monolinguals who live in Montréal and regularly overhear French in their daily life show faster learning and better retention of French voices than English monolinguals who reside in Storrs (Connecticut) and have virtually no exposure to French (Orena, Theodore, & Polka, 2015). Comparable results were reported by Perrachione & Wong (2007), who found that short-term training on a foreign language improved talker recognition in that language, but only for participants who have had previous experience with the foreign language at the outset of the experiment. More specifically, six 30-min training sessions in the lab improved identification of English speakers by Mandarin Chinese monolinguals who lived in the United States at the time, but brief training in the absence of prior exposure to Mandarin Chinese did not improve English monolinguals’ ability to identify Mandarin Chinese speakers. These findings suggest that brief exposure to an already somewhat familiar language may be adequate to re-activate a formerly existing LFE for that language, but brief exposure to an entirely unfamiliar language is not enough to cause an LFE.
Although adults need quite extensive experience with a language in a natural communicative environment before they begin to show improved talker recognition, there is reason to suspect that this may not be the case for infants. We know from previous developmental research that exposure to a new auditory or visual stimulus in the first year of life can refine infants’ perceptual abilities surprisingly quickly (Anzures et al., 2012; Fair, Flom, Jones, & Martin, 2012; Hannon & Trehub, 2005; Heron-Delaney et al., 2011; Kitamura, Panneton, & Best, 2013; Pascalis et al., 2005). For example, short-term training on a foreign language (12 sessions of 25 min live interaction with Mandarin Chinese speakers) can reverse the typical decline in foreign-language phonetic perception in English-learning 9-month-old infants (Kuhl, Tsao, & Liu, 2003), and even very brief exposure to an unfamiliar regional accent (2 min exposure to an Australian English speaker reading a familiar story) facilitates Canadian English-learning 15-month-olds’ recognition of words produced in that accent (Van Heugten & Johnson, 2014). These perceptual learning data suggest that infants may only require brief, incidental exposure to a new language or language variety for talker recognition in that language (variety) to improve.

In summary, the present study explored how language experience shapes infants’ ability to discriminate between unfamiliar speakers of that language. In Experiment 1, we tested 7.5-month-olds’ ability to distinguish between two bilingual females speaking either the infants’ native language (English) or a foreign language (Polish). In Experiment 2, we exposed English-learning 7.5-month-olds to Polish for two weeks and examined the effects of this exposure on infants’ performance in a Polish talker discrimination task. Our findings show that the LFE is robust in 7.5-month-old infants, suggesting that talker learning in infancy is reliant on language experience from the very early stages of development.
Experiment 1

In Experiment 1, English-learning infants completed a talker discrimination task in English and Polish. In alignment with our goals (to replicate Johnson et al., 2011, and to explore if language training affects voice discrimination), we tested the same age group as Johnson et al. (7- to 8-month-olds) and we used the same general testing procedure (Visual Fixation Procedure). However, we put the robustness of the LFE in 7.5-month-olds to the test by examining a different population (English monolinguals living in Canada compared to Dutch monolinguals living in the Netherlands) and language pairing (English versus Polish compared to Dutch versus Italian/Japanese). In addition, we tested infants on the voices of bilingual speakers, which ensured that any performance differences observed across language conditions were due to the change in language and not the fact that the voices in one condition were more distinct than the voices in the other condition (Winters, Levi, & Pisoni, 2008). If the LFE is robust in 7.5-month-old infants, then we should observe the same native-language advantage in talker discrimination as reported in Johnson et al. (2011). That is, English-learning infants should be able to successfully distinguish between two unfamiliar talkers when these talkers speak English but not when they speak Polish.

Method

Participants. Forty-eight full-term monolingual Canadian English-learning 7- to 8-month-olds ($M_{age} = 230$ days, range = 210–248; 26 females) from the Greater Toronto Area were tested. No hearing difficulties or recent ear infections were reported. All infants were exposed to English at least 90% of the time, and none of them had exposure to Polish or Polish-accented English prior to taking part in the study. We
chose Polish as the foreign language in this study for a variety of practical and theoretical reasons. Crucially, Polish is spoken very infrequently in the Greater Toronto Area, which helped us to ensure that infants had no ambient exposure to Polish. Fourteen additional infants were tested, but their data were excluded from the study due to failure to complete at least six habituation trials before reaching habituation criterion (10), fussing (2), failure to habituate (1), and failure to reach post-test criterion (1). The dropout rate in Experiment 1 (23%) was comparable to the dropout rates in the original Johnson et al. (2011) study (20%) and other studies using the Visual Fixation Procedure (for references see Johnson & Zamuner, 2010). Infants’ age and gender were well balanced across the English ($M_{age} = 226$ days, range = 213–244; 12 females) and Polish ($M_{age} = 233$ days, range = 210–248; 14 females) conditions.

Stimuli. Stimuli consisted of speech recordings of two English-Polish bilingual females reading 40 English and 40 Polish sentences (16–18 syllables per sentence). The English sentences were drawn from Johnson et al. (2011) and the Polish sentences were modeled after the English sentences (see Appendix A). Thirty-six sentences per language were used for habituation and the remaining four sentences were used at test (habituation and test sentences were counterbalanced across participants).

The two English-Polish bilingual speakers read all sentences in an adult-directed manner with a neutral tone of voice. Recordings (48 kHz; normalized to 69.5 dB) were made in a double-walled, sound-attenuated Industrial Acoustics Corporation (IAC) booth using high-quality recording equipment. Talker 1 was 22 years old and Talker 2 was 21 years old at time of recording. Both talkers learned English and Polish from birth and speak both languages on a regular basis. Neither of them had a Polish accent when speaking English and both of them had only very subtle English-influenced features in their Polish productions (which is to be expected considering that both were
heritage speakers of Polish who were born and grew up in the Greater Toronto Area. Both were non-smokers and neither of them had a particularly distinctive voice quality. Although we controlled for talker similarity across languages by recording bilinguals (and thus mostly eliminated the voice/language confound in previous LFE studies), we acknowledge that it is possible that even a pair of early bilingual speakers might still be more distinct in one language than another (see, e.g., Lee & Van Lancker Sidtis, 2017). However, our acoustic-phonetic analyses revealed no support for this possibility. For some acoustic-phonetic properties of the talkers’ speech productions (chosen to allow comparison with Johnson et al., 2011) see Table B.1 in Appendix B.

Procedure. Infants were tested using the Visual Fixation Procedure (Johnson & Zamuner, 2010). In this procedure, infants are presented with an auditory stimulus and a concurrent visual stimulus. As long as infants are interested in what they hear, they will look at the visual stimulus. Over time, as infants lose interest in listening, their looks to the visual display gradually decrease. Once infants’ looking time has decreased to a preset criterion (i.e., infants are habituated), a new auditory stimulus will be presented (the visual stimulus remains the same). If infants notice the change, then they will be more interested again and their looks to the visual display will increase (i.e., infants dishabituate). This sudden change in looking behavior allows us to infer that infants discriminated between the old (habituated) and new auditory stimulus. Regarding Experiment 1, we predicted that infants would dishabituate to hearing a new talker when they were tested on English, but not when they were tested on Polish.

In Experiment 1, infants sat on their caregiver’s lap in an IAC booth facing a 21.5-inch computer monitor that showed a multi-colored flickering checkerboard during all trials (see Figure 1). Loudspeakers (Alesis M1Active 520 USB) presented the speech samples at a constant, comfortable listening level. The experimenter monitored infants’
looking behavior on a separate monitor outside the booth and relayed looking data to a computer (running Habit 2000, version 2.2.4) via a keyboard. At the end of each trial, a blinking red star served to center the infant. Once the infant oriented towards the blinking star, the experimenter initiated the next trial. Caregivers wore close-fitting noise-cancelling headphones and listened to masking music intermixed with speech stimuli used in the experiment to prevent them from biasing their child’s performance.

Infants were randomly assigned to one of two language conditions: native language (English; \(N = 24\)) or foreign language (Polish; \(N = 24\)). The experiment consisted of two phases, habituation and test. During each infant-controlled habituation trial (maximum 16 s long), either two English or two Polish sentences (depending on condition) spoken by one of the talkers were repeated in a cyclic manner (inter-stimulus-interval between sentences = 300 ms; Minimum Look Time = 1 s, Minimum Look-Away Time = 2 s). The test phase began once the infant’s looking time had decreased to 65% of the initial duration (calculated over a sliding window of three trials) or the infant had completed a maximum of 18 habituation trials. Since the design of habituation and test trials was identical, caregivers and experimenter were unaware of test phase commencement. During pre-test (before the first habituation trial) and post-test (after the last test trial), the infant saw a colorful spinning windmill, which established a baseline for his/her looking behavior. If the looking time during post-test was at least 80% of the looking time during pre-test, then we would infer that the infant was still attentive to the task (and the data were therefore useable). Data for a particular infant were excluded prior to data analysis when the infant had not completed at least six habituation trials before reaching habituation criterion (which aimed to ensure that infants had gained sufficient experience with the first talker before being exposed to the second talker).
Each infant completed two ‘Same Voice’ and two ‘Different Voice’ test trials. In the Same Voice trials, the talker from the habituation phase was presented; in the Different Voice trials, the new (unfamiliar) talker was presented. Order of presentation of these two types of test trials was counterbalanced across infants (i.e., half of them heard the Same Voice trials first and half heard the Different Voice trials first). We also counterbalanced which talker was heard during habituation and which one during test. This study was approved by the University of Toronto ‘Social Sciences, Humanities and Education’ Research Ethics Board.

[INSERT FIGURE 1 ABOUT HERE]

Results and Discussion

We compared mean looking time (in seconds) during Same Voice trials to mean looking time during Different Voice trials for the two language conditions (see Figure 2). If infants noticed that a change in talker had occurred, then they should dishabituate to the new voice, i.e., their mean looking times during Same Voice and Different Voice trials should significantly differ. Based on previous habituation studies (e.g., Johnson et al., 2011; Johnson & Zamuner, 2010), we predicted that infants would look significantly longer during Different Voice than Same Voice trials in the English condition, but that their mean looking times for the two test trial types would not differ for Polish.

To assess infants’ voice discrimination abilities for English and Polish, we fit a linear mixed effects regression (LMER) model to our data using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, 2016), with mean looking time as the dependent variable and contrast-coded fixed effects for Voice
(Same, Different) and Language (English, Polish) and the Voice × Language interaction. The model also included a random intercept for subject and a random slope for Voice by subject. Model comparisons were performed to determine whether the inclusion of each fixed factor and the interaction made a significant contribution to the model. This analysis revealed a significant main effect of Voice ($\beta = 1.29$, SE $\beta = .63$, $\chi^2(1) = 4.05$, $p = .044$), suggesting that infants looked significantly longer during the Different Voice than Same Voice trials. There was no significant main effect of Language ($p = .906$), but crucially, the Voice × Language interaction was significant ($\beta = 4.26$, SE $\beta = 1.26$, $\chi^2(1) = 10.29$, $p = .001$).

To investigate the significant interaction further, we constructed separate LMER models for the two language conditions, with each model including a fixed effect for Voice, a random intercept for subject, and a by-subject random slope for Voice. In the English condition, infants looked significantly longer during Different Voice ($M = 8.8$, $SD = 4.2$) than Same Voice ($M = 5.4$, $SD = 2.9$) trials ($\beta = 3.42$, SE $\beta = 1.04$, $\chi^2(1) = 8.98$, $p = .003$). This result indicates that infants successfully detected the talker change in English. In the Polish condition, however, infants’ looking times during Different Voice ($M = 6.6$, $SD = 4.0$) and Same Voice ($M = 7.4$, $SD = 5.0$) trials did not significantly differ ($p = .282$), which suggests that infants did not notice the talker change in Polish.

[INSERT FIGURE 2 ABOUT HERE]

Taken together, the results of Experiment 1 revealed that English-learning 7.5-month-olds discriminated between two bilingual female talkers when they spoke English (a familiar language) but not when the same two females spoke Polish (an
unfamiliar language). This result aligns well with the pattern of results reported in Johnson et al. (2011) for Dutch-learning 7.5-month-olds, who showed better talker discrimination for Dutch than an unfamiliar language. In both studies, infants detected a talker change in their native language but not in a non-native language. This finding clearly demonstrates that Johnson et al.’s findings generalize to a different population, language pairing, and the use of the voices of bilinguals rather than monolinguals.

Interestingly, previous developmental research on phonetic perception has shown that the decline in foreign-language phonetic perception can be reversed by exposing infants to non-native phonetic contrasts (Kuhl et al., 2003). Hence, is it possible that foreign-language exposure improves foreign-language talker recognition in a similar manner? And if so, how much and what type of exposure to the foreign language is necessary to facilitate talker recognition for that language? To address these questions, Experiment 2 tested whether short-term familiarization with Polish would improve English-learning infants’ ability to tell apart Polish speakers. If brief exposure to a new language is enough for infants to learn about the sound structure of that language, and if the LFE is indeed driven by language language-specific phonology, then infants tested in Experiment 2 (unlike infants tested in Experiment 1) should be able to tell apart the bilingual talkers even when they speak Polish. However, if infants (like adults) require more extensive exposure to a foreign language for sufficient phonological learning to take place, then infants in Experiment 2 should perform similarly to infants in Experiment 1 (Polish condition).

**Experiment 2**

In Experiment 1, English-learning 7.5-month-olds readily distinguished between the voices of English-Polish bilinguals when they spoke English but not when they
spoke Polish, thus clearly demonstrating an LFE for English. Experiment 2 examined whether we could experimentally improve infants’ talker discrimination skills for Polish speakers by familiarizing them with the Polish language before completing a talker discrimination experiment. This experiment was a repetition of the Polish condition of Experiment 1, with the only difference that infants in Experiment 2 were exposed to a 15-min audio recording of Polish speakers at home for two weeks daily before participating (note that these speakers were different from those recorded for the discrimination experiment). We exposed infants to multiple speakers during training to provide them with a diverse sample of the new language (see, e.g., Perrachione & Wong, 2007) and because talker variability has been shown to promote language learning in infants (Houston & Jusczyk, 2000; Rost & McMurray, 2009). Testing infants at 7.5 months of age was considered appropriate for this training study because it allowed us to directly compare infants’ performance in Experiments 1 and 2, and because at 7.5 months, infants still refine their perceptual abilities quickly and efficiently (see Introduction for references, and in particular Kuhl et al., 2003). If 7.5-month-old infants are equipped to rapidly acquire basic phonological knowledge of a language from brief auditory exposure, then infants in Experiment 2 should show better talker discrimination for Polish than infants in Experiment 1 (Polish condition). However, if infants (like adults) need extensive language experience for talker learning in that language to improve, then infants’ performance should not significantly differ between experiments.

To provide an adult comparison (and to facilitate comparison to previous adult studies), we tested infants’ mothers on a ‘voice line-up’ talker recognition task in Polish. Since these mothers had played the Polish recording for their child, they too had been familiarized with Polish at home for two weeks. To examine whether these
mothers performed better than mothers without previous experience with Polish, we additionally tested a control group of mothers who visited our lab to participate with their child in an unrelated study.

Method

Participants. Forty-four full-term monolingual Canadian English-learning 7- to 8-month-olds ($M_{age} = 229$ days, range = 212–245; 24 females) from the Greater Toronto Area were tested (eligibility criteria were the same as in Experiment 1). Like in Experiment 1, infants were carefully screened prior to participating to ensure that they had no exposure to Polish in their everyday life. The data from an additional 25 infants were excluded due to failure to complete at least six habituation trials before reaching habituation criterion (15), fussing (4), failure to habituate (4) or meet the language requirements (1), and one caregiver not following the home-exposure instructions (36% dropout rate). Note that we tested a large sample of infants in Experiment 2 to increase power (we expected that if an effect of language exposure on voice discrimination occurred, then the effect would still be relatively small). We also tested 36 native English-speaking mothers of infants who participated in Experiment 2 ($M_{age} = 34$ years, $SD = 5$) and 36 native English-speaking mothers of children who participated in a different study in our lab ($M_{age} = 34$ years, $SD = 4$).

Stimuli. The auditory stimulus used for the at-home exposure stage of Experiment 2 consisted of an audio recording of three male and three female native Polish speakers (all professional actors and actresses, ranging in age from early thirties to late sixties). The recording was a compilation of passages extracted from various audiobooks (e.g., Cinderella; The Ugly Duckling), which were made freely available online by the Modern Poland Foundation (Wolne Lektury project). To keep infants’
attention while listening to the stories, we randomized several shorter passages of each speaker on the recording (rather than presenting each speaker in a block) and included short musical interludes and sound effects (e.g., bells, chimes, toy rattles). The recording (48 kHz; normalized to 69.5 dB) was 15 min long.

**Procedure.** The procedure of the first stage of Experiment 2, during which infants were familiarized with Polish, was as follows: Caregivers of participating infants were instructed to play the recording of Polish speakers to their child at home once every day for the two weeks preceding their scheduled appointment in the lab. They were informed that they could play the recording at any time during the day as long as their child was awake and alert. They were asked to avoid background noise (e.g., from television or radio) while playing the recording. In addition, caregivers completed a diary in which they kept records of the date and time the recording was played to their child, the playback device (most commonly a smartphone), their child’s attention level (on a scale from 1 ‘not attentive’ to 7 ‘very attentive’), and any additional comments (e.g., what the child was doing while listening). Caregivers were unaware that they would be tested on a Polish talker recognition experiment until they visited the lab after the 2-week exposure stage.

The procedure of the second stage of the experiment, during which infants visited our lab and completed a talker discrimination task in Polish, was identical in every respect to the procedure of the Polish condition of Experiment 1.

During their lab visit, the two groups of mothers (i.e., mothers who participated with their child in Experiment 2 and thus had been exposed to Polish for two weeks, and mothers who participated with their child in an unrelated study and had no familiarity with Polish) were tested on a talker recognition experiment in Polish. Since the habituation paradigm used to test infants in this study is not appropriate for testing
adults, we tested mothers on a ‘voice line-up’ procedure using the same stimuli that we used for testing infants (along with stimuli produced by two additional Polish-English bilinguals). Each trial of the experiment (four trials in total) consisted of three stages. In Stage 1, mothers were familiarized with one of the talkers (reading four Polish sentences). In Stage 2, mothers were engaged in a 1 min long movie showing episodes from children’s TV shows and movies (dubbed with music and sound effects but no speech). In Stage 3, mothers were presented with the talker from Stage 1 and a different talker, and were asked to select which of the two talkers was the talker from Stage 1.

Results and Discussion

Experiment 2 tested whether infants who had been exposed to stories read in Polish for two weeks daily would show better talker discrimination than infants who participated in the Polish condition of Experiment 1 and who were entirely unfamiliar with Polish. To allow for a statistical comparison between the two groups of infants, we submitted the looking time data from Experiment 2 alongside the looking time data from the Polish condition of Experiment 1 to an LMER model. This model included contrast-coded fixed effects for Voice (Same, Different) and Exposure (Polish pre-exposure, No Polish pre-exposure) and the interaction, as well as a random intercept for subject and a by-subject random slope for Voice. No significant main effects of Voice and Exposure (and no interaction) were found ($\chi^2 < .464$, $p > .445$). This indicates that irrespective of Exposure condition, mean looking times for Same Voice and Different Voice trials were not significantly different (i.e., infants did not detect the talker change in either experiment; see Figure 3). Note that infants’ performance in Experiment 2 was not significantly correlated with the attention rating provided by the caregivers ($M = 4.2$, $SD = 1.1$; $r = -.07$, $p = .691$).
We then examined mean recognition accuracy for the two groups of mothers who participated in the ‘voice line-up’ experiment, and found that performance of the group with Polish experience ($M = .81, SD = .18$) and the group without Polish experience ($M = .75, SD = .21$) did not significantly differ ($p = .229$). This finding is consistent with adult data from previous studies (especially Perrachione & Wong, 2007) and confirms that a total of 3–4 hrs exposure to a foreign language does not enhance adults’ foreign-language talker recognition abilities.

In summary, Experiment 2 showed that infants failed to notice a talker change in Polish even after they had been familiarized with Polish. This demonstrates that daily at-home exposure to audio recordings of Polish speakers for 14 consecutive days proved insufficient to increase infants’ sensitivity to talker differences in Polish (and note that this exposure was also not enough to increase mothers’ talker recognition performance in Experiment 2; see Perrachione & Wong, 2007). While 5 hrs exposure to a foreign language has been shown to influence foreign-language phonetic perception in infants (Kuhl et al., 2003), 3.5 hrs foreign-language training in the current study did not suffice to facilitate foreign-language talker recognition (i.e., neither infants nor adults developed an LFE for the foreign language after short-term familiarization with the language).

**General Discussion**

Understanding the developmental origins of the LFE is not only important to explaining the LFE, it is also important to understanding when infants become adult-
like in their processing of indexical and linguistic information present in the speech signal. Past studies have explored the parameters driving the LFE by testing listener populations whose phonological knowledge of a language is compromised to one degree or another, such as late bilinguals and second-language learners (Bregman & Creel, 2014; Goggin et al., 1991; Köster & Schiller, 1997), dyslexic listeners (Perea et al., 2014; Perrachione et al., 2011), or school-aged children (Levi & Schwartz, 2013; Levi, 2017; Perea et al., 2014). But to date, only one published study has examined the LFE in infants (Johnson et al., 2011). In the current study, we show that the LFE appears to be robust in the second half of the first year of life. Like Johnson et al. (2011), we found that infants were more sensitive to fine-grained phonetic differences in the voices of native-language speakers than in the voices of foreign-language speakers, thus providing clear evidence of the LFE in 7.5-month-olds.

The finding of a robust infant LFE in Johnson et al. (2011) and the present study makes an important contribution to the ongoing discussion in both the developmental and adult literature about the linguistic factors that are driving the LFE. Since we know that infants’ comprehension skills are highly restricted at 7.5 months, our findings lend additional support to the idea that the LFE is predominantly caused by reliance on phonology rather than access to lexical/semantic information (e.g., Bregman & Creel, 2014; Johnson et al., 2017; Kadam et al., 2016; Levi & Schwartz, 2013; Orena et al., 2015; Perrachione et al., 2011).

While much evidence points towards a central role of phonology in explaining the LFE, so far only few attempts have been made, firstly, to narrow down the specific aspects of phonology that are causing the effect (e.g., Schiller, Köster, & Duckworth, 1997), and secondly, to elaborate on why phonological knowledge of the familiar language would account for the performance difference between talker recognition in
the familiar versus an unfamiliar language. By firmly establishing that infants as young as 7.5 months show the LFE, the current study suggests that the LFE can at least minimally be driven by language-specific knowledge that infants have already acquired at 7–8 months. At this age, infants possess some knowledge of the prosodic and segmental structure of their native language (Johnson, 2016; Jusczyk, 1997). For example, they recognize the rhythm of their native language (Nazzi, Jusczyk, & Johnson, 2000), know where lexical or phrasal stress typically falls in words (Jusczyk, Houston, & Newsome, 1999), prefer familiar over unfamiliar language varieties (Kitamura et al., 2013), have started to acquire how tones and phones are used to distinguish between word forms (Mattock, Molnar, Polka, & Burnham, 2008), and are highly attuned to the distribution of sounds and syllables in their native language (Ngon et al., 2013; Romberg & Saffran, 2010). This type of emerging phonological knowledge must be enough to elicit an LFE for talker recognition. To further pinpoint what aspects of phonology are responsible for the emergence of the LFE, future research needs to establish a developmental time course for the LFE by testing whether younger infants also show the LFE.

But why does familiarity with a language’s phonology play such an central role in explaining the LFE? How does phonological knowledge contribute to establishing talker identity from native- versus foreign-language speech? To determine who is speaking, listeners need to analyze the incoming speech signal with respect to various linguistic and indexical properties. Some of the parameters that most effectively signal talker identity are the intensity, spectral and temporal characteristics of individual phonemes (along with coarticulatory and connected-speech processes). Having access to this type of information (for example, when listening to native-language speech) allows listeners to compare different speakers’ phonetic realisations of phonemes (allophones)
and to establish similarities or differences between speakers (thus facilitating successful native-language talker recognition). However, if language-specific allophonic cues to talker identity are not as readily available to the listener (for example, when listening to foreign-language speech), less information will be available on which to base a decision about talker identity (and as a result, talker recognition will be compromised). Hence, the emergence of the LFE may be explained, at least in part, by reference to these differences in processing talker-specific phonological information in native- versus foreign-language speech.

In the present study, we not only established that the LFE is robust in 7.5-month-olds, we also began examining the specific learning conditions under which the LFE can be elicited in infants. Motivated by findings in the perceptual narrowing literature (e.g., Anzures et al., 2012; Fair et al., 2012; Heron-Delaney et al., 2011; Pascalis et al., 2005), we tested whether short-term exposure to a new language would be enough for infants to reach the level of linguistic proficiency needed to improve talker recognition in that language. We predicted that infants and adults might differ in the amount of exposure needed to cause the LFE (specifically, that infants might need less exposure than adults). Our study showed that two weeks of short daily exposure to Polish audio recordings had no measurable effect on infants’ performance in a Polish talker discrimination task, which could be taken as evidence that infants are not set up to develop an LFE faster than adults. But can we be sure that the limited amount of language exposure that infants received in our study was the reason that we did not observe an effect of exposure in Experiment 2?

One reason that infants did not show improved talker discrimination for Polish after Polish training could certainly be that infants (like adults) require more substantial experience with a foreign language before talker recognition in the foreign language
improves. However, another possibility is that the learning conditions in Experiment 2 were not favourable to infants, and specifically, that the type of language exposure that infants received (audio recordings of Polish speakers) did not support acquisition of the sound structure of the new language. Exposing infants live to foreign-language speakers in a more social communicative environment (as per Kuhl et al., 2003) might have increased infants’ attention (or motivation to learn about the new language) and helped them to more efficiently extract relevant phonological properties of the foreign language. Moreover, perhaps infants did not exhibit an LFE for Polish in Experiment 2 because they were already too old to attune to the phonological structure of a new language from brief auditory language exposure alone. As infants grow up and their developing perceptual system becomes more specialized through a process referred to as perceptual narrowing (Watson, Robbins, & Best, 2014), more experience with a foreign language may be needed for sufficient language learning to occur. That said, we currently do not know whether our results were indeed caused by lack of sound structure acquisition during the 14-day training period, or in fact by difficulty with applying this newly-acquired knowledge to a different task situation (talker identification). Hence, future research should develop a paradigm that elicits improvements in phonological knowledge and tests whether this new knowledge directly transfers to benefits in talker identification.

In conclusion, the present study makes an important contribution to our understanding of the developmental changes in talker recognition brought about by linguistic experience with a first or second language. Firstly, we established that the native-language advantage for talker identification is robust in 7.5-month-old infants, suggesting that the Language Familiarity Effect is likely rooted in early-emerging knowledge of native-language phonology. Secondly, we explored the learning
conditions that promote language-specific talker learning in infancy, and concluded that both infants and adults may require substantial (potentially live) exposure to a foreign language before they begin to show evidence of the LFE. Taken together, our findings not only benefit the community of researchers studying the linguistic and perceptual mechanisms underlying the LFE in adults, they also provide a window into when children’s processing of linguistic and indexical information becomes adult-like in this manner and thus inform developmental models of talker recognition and spoken language processing more generally. Hence, our work lays a solid foundation for future work examining how talker recognition and spoken language processing are related across languages and ages, how this relationship evolves from infancy to adulthood, and what this means for the development of talker recognition and language learning across the lifespan more generally.
Acknowledgements

The authors thank Lisa Rustom, Natalie Rzeszutek and Lisa Hotson for assistance with stimulus creation and data collection, Angela Cooper for assistance with data analysis, and all participating families. This research was supported by grants awarded to EKJ from the Social Sciences and Humanities Research Council of Canada, the Natural Sciences and Engineering Research Council of Canada, and the Canada Research Chairs Program.
Appendix A

Examples of English test sentences

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

Examples of Polish test sentences

1. Pociąg odjechał ze stacji więcej niż piętnaście minut temu. (The train left the station more than fifteen 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.)

Appendix B

[INSERT TABLE B.1 ABOUT HERE]
References


THE LANGUAGE FAMILIARITY EFFECT IN INFANTS


Ngon, C., Martin, A., Dupoux, E., Cabrol, D., Dutat, M., & Peperkamp, S. (2013). (Non)words, (non)words, (non)words: evidence for a protolexicon during the first


R Core Team (2016). R: A language and environment for statistical computing. R


Figures

Figure 1. Using the Visual Fixation Procedure, 7.5-month-old infants were shown a high-contrast visual stimulus (checkerboard) while listening to the voices of English-Polish bilingual speakers. During habituation, infants were familiarized with the voice of a young woman speaking either in English or in Polish. During test, infants were tested on their ability to detect a talker change (within languages).
Figure 2. Mean looking times during Same Voice and Different Voice test trials for English (native language) and Polish (foreign language) in Experiment 1 (error bars indicate SE). Looking time difference between test trial types was statistically significant for English but not for Polish (indicative of a Language Familiarity Effect).
Figure 3. Mean looking times during Same Voice and Different Voice test trials in Experiment 2 (Polish pre-exposure) and for comparison, in the Polish condition of Experiment 1 (No Polish pre-exposure). Like infants who were not given pre-exposure to Polish in Experiment 1, infants in Experiment 2 (who were familiarized with Polish before completing the talker discrimination task) did not notice a talker change in Polish (error bars indicate SE).
Table B.1. Acoustic-phonetic measurements for two female English-Polish bilingual speakers (SDs are given in parentheses).

<table>
<thead>
<tr>
<th>Language Condition</th>
<th>Mean F0 (Hz)</th>
<th>Standard Dev. F0 (Hz)</th>
<th>Sentence Length (s)</th>
<th>Articulation Rate (syll./s)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Talker 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>200.2 (8.3)</td>
<td>57.5 (10.0)</td>
<td>4.2 (0.4)</td>
<td>4.1 (0.3)</td>
</tr>
<tr>
<td>Polish</td>
<td>210.9 (9.2)</td>
<td>61.8 (10.4)</td>
<td>4.4 (0.3)</td>
<td>3.9 (0.4)</td>
</tr>
<tr>
<td><strong>Talker 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>210.1 (9.1)</td>
<td>56.2 (12.8)</td>
<td>3.7 (0.3)</td>
<td>4.7 (0.4)</td>
</tr>
<tr>
<td>Polish</td>
<td>205.1 (9.5)</td>
<td>56.3 (16.3)</td>
<td>4.6 (0.5)</td>
<td>3.8 (0.4)</td>
</tr>
</tbody>
</table>