Effects of language experience and task demands on talker recognition by children and adults[a)]

Natalie Fecher[b)] and Elizabeth K. Johnson
Department of Psychology, University of Toronto, 3359 Mississauga Road, Mississauga, Ontario, L5L 1C6, Canada

(Received 11 December 2017; revised 24 March 2018; accepted 30 March 2018; published online 25 April 2018)

Talker recognition is a language-dependent process, with listeners recognizing talkers better when the talkers speak a familiar versus an unfamiliar language. This language familiarity effect (LFE) is firmly established in adults, but its developmental trajectory in children is not well understood. Some evidence suggests that the effect already exists in infancy, but little is known about how it unfolds in childhood. The present study explored whether the strength of the LFE increases in early childhood. Adults and children were tested in their native language and a foreign language using a "same-different" talker discrimination task and a "voice line-up" talker recognition task. Results showed that adults and 6-year-olds, but not 5-year-olds, exhibit a robust LFE, suggesting that the effect strengthens as children's language competence increases. For both adults and older children, the emergence of an LFE moreover appeared to be task-dependent. This study contributes to a better understanding of how children develop mature talker recognition abilities and when children's processing of indexical and linguistic information in speech approaches adult-like levels. Furthermore, the findings reported here contribute to the debates regarding the origins of the LFE—a hallmark of adult talker recognition. © *2018 Acoustical Society of America*.
https://doi.org/10.1121/1.5032199

[MS] Pages: 2409–2418

I. INTRODUCTION

The linguistic mechanisms underlying talker recognition and spoken language processing have traditionally been studied independently of one another. Classic models of speech perception proposed that talker-specific information in the speech signal is normalized before listeners can access the abstract representations of speech sounds (e.g., Goldinger, 1996; Johnson, 1990). However, more recent evidence suggests that listeners process both types of information in an integrated manner. For example, word recognition is facilitated when the words are spoken by a familiar talker (e.g., Mullennix *et al.*, 1989; Nygaard and Pisoni, 1998); talker recognition is dependent on the segmental content of an utterance (e.g., Andics *et al.*, 2007; Cutler *et al.*, 2011); and in what has been termed the language familiarity effect (LFE), adults recognize talkers who speak a familiar language better than talkers who speak an unfamiliar language. In the current study, we seek to better characterize the emergence of mature talker recognition abilities in children by examining the development of the LFE—a core property of adult talker recognition—in young children and adults. We hypothesize that as talker recognition and spoken language processing do not reach maturity until late childhood, so too will the LFE show a protracted period of development.

In one of the earliest studies of the LFE, Thompson (1987) reported that monolingual English listeners identified the voices of bilingual talkers better when these talkers spoke English than when the same talkers spoke Spanish. Some researchers have posited that lexical knowledge accounts, at least in part, for the performance difference between native and non-native talker recognition in adults (e.g., Goggin *et al.*, 1991; Perrachione *et al.*, 2015). However, there is growing evidence that knowledge of native-language sound structure may be the locus of the LFE, with studies suggesting that the LFE is caused by listeners' reliance on familiar phonology to identify talkers (e.g., Fleming *et al.*, 2014; Johnson *et al.*, 2018). The latter argument is supported by research showing that level of reading proficiency (which is linked to phonological processing abilities) influences native- and foreign-language talker recognition in dyslexic (Perrachione *et al.*, 2011) and non-dyslexic (Kadam *et al.*, 2016) adults (but see Perea *et al.*, 2014), and that degree of bilingualism (which is presumably linked to degree of mastery of phonology) has a gradient effect on talker learning (Bregman and Creel, 2014; Köster and Schiller, 1997; Orena *et al.*, 2015; Sullivan and Schlichting, 2000). Perhaps the strongest evidence in support of the role of phonology in explaining the LFE is that even infants with highly limited comprehension skills show the LFE in a talker discrimination task (i.e., 7.5-month-old infants robustly tell apart speakers of their native language but not speakers of a foreign language; Fecher and Johnson, 2018; Johnson *et al.*, 2011).

The finding that infants exhibit an LFE raises the question of how the LFE unfolds in later child development. If even infants already show the effect, then does this mean that the native-language benefit for talker recognition

[a)]Portions of this work were presented at the 15th Conference on Laboratory Phonology, Ithaca, NY, USA, July 2016.
[b)]Electronic mail: natalie.fecher@utoronto.ca

reaches full adult-like strength in infancy? The existence of an LFE at only 7.5 months of age could be taken as evidence for this view. Alternatively, the LFE, like many other aspects of spoken language processing, might show a protracted period of development and not reach adult-like levels until much later in childhood. We know from independent research on speech processing in young children that both talker recognition (e.g., Bartholomeus, 1973; Creel and Jiménez, 2012; Levi, 2017; Mann *et al.*, 1979) and spoken language processing (e.g., Hazan and Barrett, 2000; Nittrouer, 1992; Ohde and Haley, 1997) continue to improve throughout childhood. If the maturation of these two abilities were related in development, then we would predict that talker recognition in children's native language (but not necessarily in a foreign language) improves as children become more proficient in their native language; as a consequence, the LFE would grow stronger over time. Is there any indication in the existing literature of which of the two proposed developmental pathways—the LFE being fully developed in infancy versus the strength of the LFE increasing throughout childhood—best reflects the development of the native-language benefit for talker recognition?

A closer look at previous work on the LFE reveals that our current understanding of the developmental time course of the LFE is very limited. The few studies that have previously tested children on the LFE were not designed to compare LFE strength across different ages, and the findings from these studies have been mixed. Levi (2017) tested native- and foreign-language talker discrimination in 6- to 11-year-old mono- and bilingual children (bilinguals performed better overall), Levi and Schwartz (2013) tested 7- to 12-year-old children with typical language development or specific language impairment on talker discrimination in their native language and a foreign language, and Perea *et al.* (2014) tested native- and foreign-language talker learning in 7- to 15-year-old children with or without dyslexia (neither Levi and Schwartz, 2013, nor Perea *et al.*, 2014 found a significant difference between the two groups of children that they tested). While Levi and Schwartz (2013) report a robust LFE for adults and a weak LFE for the younger (non-impaired) children (suggesting that the LFE might indeed grow stronger over time), Perea *et al.* (2014) and Levi (2017) found no difference in LFE strength across development. Hence, the question of whether the LFE is equally strong in children and adults, or whether it strengthens between childhood and adulthood, remains open.

Since the published literature on the LFE fails to paint a clear picture regarding when, how, and why the LFE develops in children, the present study begins to address these issues by asking whether the strength of the LFE differs between children and adults. Here, we put our two alternative hypotheses—the LFE is fully developed in infancy versus the LFE grows stronger in childhood—to the test by examining adults' and children's talker recognition abilities in a familiar language compared to an unfamiliar language. Establishing the developmental trajectory of the LFE can help us to better understand when children's processing of linguistic and indexical information in speech approaches adult-like levels. It may also provide an explanation for why it takes children so long to reach mature talker recognition abilities, and thus advance our understanding of the processes underlying talker recognition more generally. Finally, studying the developmental roots of the LFE contributes to the debates about whether the LFE is driven by higher-level linguistic or lower-level acoustic-phonetic information in speech.

To summarize, the current study examined how adults and children process talker identity information in native- and foreign-language speech. In Experiment 1, we used an AX "same-different" talker discrimination task to compare native English-speaking adults' and 5- and 6-year-old children's ability to tell apart the voices of English–Polish bilinguals speaking either a familiar language (English) or an unfamiliar language (Polish). In Experiments 2 and 3, adults and children were tested on talker recognition in a familiar versus an unfamiliar language using a child-friendly "voice line-up" procedure. If the LFE is fully matured by the time children reach 5 to 6 years of age, then we would predict that LFE strength does not differ between children and adults. However, if the LFE strengthens in childhood (because both talker recognition and spoken language processing are still improving), then we would predict that the LFE is weaker in children than adults (because children do not yet have adult-like native-language competence). Based on our findings, we conclude that (1) the LFE continues to mature with age even past infancy, and (2) the emergence of the LFE may crucially depend on the task and memory demands of the experimental tasks used to test talker identity processing in adults and children.

II. EXPERIMENT 1

In Experiment 1, we used an AX "same-different" talker discrimination task to test whether the LFE is stronger in children than adults despite existing evidence that a native-language benefit for talker discrimination is already present in infants. Adults and 5- and 6-year-old children were presented with the voices of bilingual English–Polish speakers. Bilingual speakers were recorded to reduce the likelihood that any differences in discrimination accuracy between language conditions resulted from the language change and not from a voice change (see, e.g., Levi, 2017; Winters *et al.*, 2008). We tested children at 5 and 6 years of age because, first, these children were younger than the children tested in previous studies of the LFE (Levi, 2017; Levi and Schwartz, 2013; Perea *et al.*, 2014); second, children's talker recognition and spoken language processing skills are still improving at this age (making this age group particularly suitable for studying the relationship between these two domains); and third, we aimed to get one step closer to providing an explanation for why it takes children until early adolescence to reach adult-like talker recognition abilities (and previous research suggests that this process takes at least until age ten; e.g., Mann *et al.*, 1979). We used a simple AX discrimination task to maximize the odds that children would succeed in our task (i.e., to avoid a floor effect sometimes observed in studies on talker identity processing in young children; e.g., Creel and Jiménez, 2012). We predicted that

TABLE I. Acoustic-phonetic measurements for four female English–Polish bilingual speakers. Data in parentheses are standard deviations. Hz = Hertz; s = second; syll. = syllable.

Talker	Language condition	F0 (Hz) *M*	F0 (Hz) *SD*	Sentence length (s)	Articulation rate (syll./s)
Talker 1	English	194.4 (8.4)	42.8 (11.1)	4.0 (0.3)	4.3 (0.3)
	Polish	206.6 (4.9)	51.3 (3.2)	4.3 (0.3)	3.8 (0.2)
Talker 2	English	205.0 (2.8)	33.9 (7.1)	3.5 (0.4)	4.8 (0.5)
	Polish	195.3 (10.1)	38.6 (8.4)	4.7 (0.3)	3.7 (0.2)
Talker 3	English	180.1 (7.1)	45.4 (9.9)	3.3 (0.2)	5.2 (0.4)
	Polish	191.5 (8.0)	39.2 (9.0)	3.6 (0.2)	4.7 (0.3)
Talker 4	English	153.9 (12.8)	41.1 (9.1)	3.6 (0.2)	4.7 (0.4)
	Polish	178.8 (7.0)	35.4 (8.4)	4.8 (0.5)	3.7 (0.4)

adults would perform better overall than children. In addition, we predicted that if the LFE is fully developed by 5 to 6 years of age, then LFE strength should not differ between children and adults; however, if the LFE continues to mature during childhood, then the LFE should be weaker in children than adults.

A. Methods

1. Participants

Thirty-two native English-speaking adults [$M_{age} = 20.3$ years, standard deviation (SD) = 1.6; 26 female] and 32 5- and 6-year-old children ($M_{age} = 6.0$ years, SD = 0.3; 14 female) from the Greater Toronto Area were tested. Thirteen additional children were tested, but their data were excluded due to a response bias (i.e., these children consistently gave the same response in at least 90% of trials; ten participants), unwillingness to complete the study as instructed (2), and technical error (1).

Adults were native speakers of English who learned English from birth in North America and whose dominant language is English. Children were monolingual North American English speakers who were exposed to English at least 90% of the time. Both adults and children were unfamiliar with Polish prior to participating. None of the participants (or their caregivers) reported a history of ear infections or speech, language, and hearing disorders, and all had normal or corrected-to-normal vision. Adults received course credit or $10 compensation for participating, and children received a small gift. In all three reported experiments, adults or children's caregivers provided written consent for their own or their child's participation, and children verbally assented to taking part in the study.

2. Stimuli

Stimuli consisted of speech recordings of four female English–Polish bilingual speakers 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).

The sentences were read in an adult-directed manner with a neutral tone of voice. Recordings were made in a double-walled, sound-attenuated Industrial Acoustics Company (IAC) booth using high-quality recording equipment, and stored in WAVE Audio File Format (48 kHz; normalized to 69.5 dB). The talkers' mean age at time of recording was 21.3 years (SD = 0.5). All four talkers learned English in North America, learned both English and Polish before the age of ten, and speak both languages on a regular basis (for some basic acoustic-phonetic measurements of the talkers' speech productions see Table I). We can rule out any effects of foreign accent experience on talker recognition performance in our study, because in contrast to Levi (2017), where the speakers had a non-native accent in the native-language condition (English with a German accent), none of the speakers in our study had a Polish accent when speaking English.

Visual stimuli consisted of cartoon aliens of various shapes and colors (see Fig. 1 and Fig. 3). The alien characters were randomly assigned to the voices across experimental trials because the task was not to learn to associate a character with a voice, but to decide (on each trial) whether the voices sounded like they belonged to the same talker or two different talkers. The pairs of characters shown on any given trial matched in shape but differed in color in order to reduce the likelihood that children would show a preference for one character over another. Each participant saw the same character only once during the experiment.

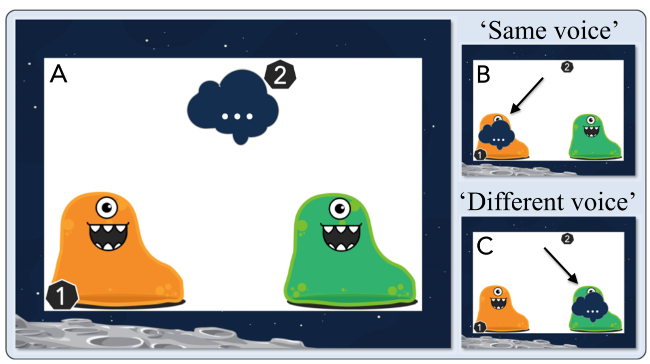

FIG. 1. (Color online) Sample experimental trial in Experiment 1. On each trial, participants clicked the alien labelled "1" to hear the first sentence. Then they clicked the speech cloud (labelled "2") to hear the second sentence (see panel A). After that, they gave a "same voice" or "different voice" response by dragging the speech cloud onto the appropriate alien. If they thought the second sentence was spoken by the alien who spoke the first sentence, then they dragged the speech cloud onto that alien (see panel B). If they thought the second sentence was *not* spoken by the alien that they had heard before, then they dragged the speech cloud onto the other alien (which always differed in color but not in shape; see panel C).

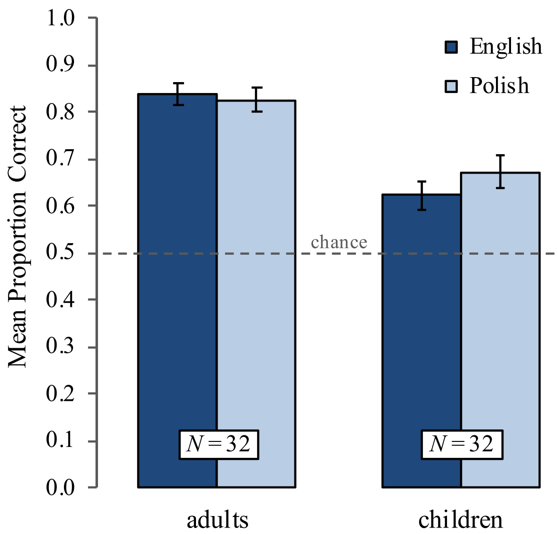

FIG. 2. (Color online) Mean proportion correct for English and Polish in Experiment 1. Neither age group showed evidence of an LFE on talker discrimination (error bars indicate SE).

3. Procedure

Participants were tested individually in a quiet testing room during a single experimental session. The experiment was designed in Articulate Storyline 2. Stimulus presentation and data collection were controlled online using the learning management system CourseSites by Blackboard. Visual stimuli were presented on a 15-in. touch-sensitive computer monitor (ELO Touchsystems 1725L). Auditory stimuli were presented binaurally over Sennheiser HD 280 PRO headphones at a constant, comfortable listening level.

The procedure in Experiment 1 was the same for adults and children. On each trial of the AX talker discrimination task, participants heard two sentences spoken either both in English or both in Polish, and were asked to judge whether these sentences had been produced by the same talker or two different talkers. Sentences were never repeated within or across trials, which prevented listeners from associating any given sentence with a specific talker and required them to generalize talker-specific information across sentences.

Upon arrival in the lab, participants were told that they would play a computer game, and that in each round of the game they would see two aliens and a "speech cloud" floating midway above the two aliens. Participants were instructed to first click the alien labelled "1" to listen to a sentence, and after that click the speech cloud (labelled "2") to listen to another sentence (see Fig. 1, panel A). They could listen to each sentence only once. Their task was to decide whether the second sentence was spoken by the alien that they had just heard ("same voice" response) or by a different alien ("different voice" response). Participants made their choice by dragging the speech cloud onto the appropriate alien (see Fig. 1, panels B and C). They did not receive feedback on the correctness of their responses.

To ensure that participants understood the task, the experimental trials were preceded by three practice trials in English (consisting of acoustically-distinct voices that were not included in the experimental trials). After completion of the practice trials, participants were informed that they would hear a foreign language in some of the forthcoming trials, but they were not told what the foreign language would be (as per Levi, 2017; Levi and Schwartz, 2013; Winters *et al.*, 2008). If a participant consistently responded either "same voice" or "different voice" in at least 90% of experimental trials (indicative of a response bias), the data for this participant were excluded prior to data analysis. The experiment was self-paced; adults took around 5 min to complete the experiment and children typically needed a few more minutes.

The experiment consisted of 20 trials (ten English, ten Polish) presented in random order. Each voice was presented equally often per language condition and each voice was combined with every other voice. We counterbalanced the order of presentation of talkers (i.e., which talker was heard first and which talker was heard second), and the side of the alien who spoke the first sentence (i.e., half of the participants always saw this alien on the left and the other half always saw this alien on the right side of the screen).

B. Results and discussion

We analyzed mean proportion correct talker discrimination (see Fig. 2). One-sample *t*-tests showed that both adults ($M_{English} = 0.84$, SD = 0.12; $M_{Polish} = 0.83$, SD = 0.15) and children ($M_{English} = 0.62$, SD = 0.17; $M_{Polish} = 0.67$, SD = 0.19) distinguished between the talkers at levels significantly above chance (0.5) regardless of language condition (all $p < 0.001$).

To assess adults' and children's talker discrimination performance, we fit a generalized mixed-effects model to our data using the *glmer* function of the *lme4* package (Bates *et al.*, 2015) in *R* (R Core Team, 2016). The contrast-coded independent variables age (0.5: adults, −0.5: children) and language (0.5: English, −0.5: Polish) and the age × language interaction were entered as fixed effects. We also included random intercepts for participant and talker, and a random slope for language by participant. The model revealed a significant main effect of age [$\beta = 1.04$, standard error (SE) = 0.19, $z = 5.53$, $p < 0.001$], with adults performing better than children (irrespective of language condition). However, the model revealed no significant main effect of language and no significant age × language interaction ($p > 0.313$), suggesting that performance did not differ between English and Polish (for either age group).[1]

As predicted, adults outperformed children. Nevertheless, the children performed better than we expected based on previous work with young children (above chance for both languages; see, e.g., Bartholomeus, 1973; Creel and Jiménez, 2012; Levi, 2017, Levi and Schwartz, 2013; Mann *et al.*, 1979; Perea *et al.*, 2014; Spence *et al.*, 2002; Van Heugten *et al.*, 2014). Note, however, that a comparison of developmental studies on talker identity processing is generally difficult due to differences in stimuli (e.g., words versus sentences), experimental paradigms (e.g., talker learning, voice line-up, or discrimination tasks), degrees of familiarity with the talkers (e.g., classmates or well-known cartoon characters versus unfamiliar talkers), exclusion criteria (e.g., response bias thresholds), or age means and ranges.

Contrary to our predictions, neither adults nor children showed evidence of a native-language benefit for talker

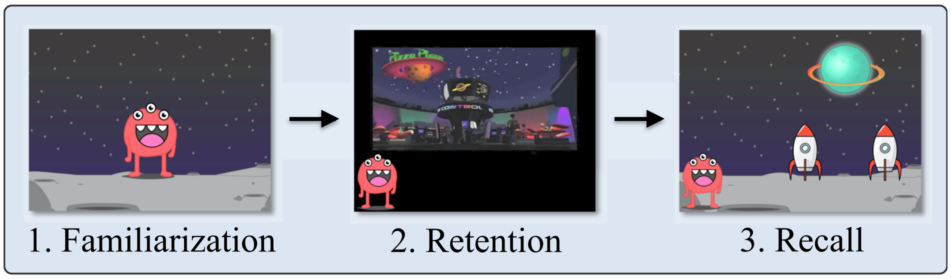

FIG. 3. (Color online) Sample experimental trial in Experiment 2. Each trial consisted of three stages: (1) familiarization, where participants were exposed to the target voice, (2) retention, where participants watched a short cartoon movie, and (3) recall, where participants identified the target voice in a voice line-up consisting of the target and a distractor voice.

discrimination. Considering the mixed results in previous child studies on the LFE (Levi, 2017; Levi and Schwartz, 2013; Perea *et al.*, 2014), the absence of an LFE for the children was less surprising. For the adults, however, this result was surprising because the LFE is considered to be a robust perceptual phenomenon that has been replicated across various listener populations and language pairings (see Perrachione, 2017).

The absence of an LFE in Experiment 1 may be attributed to the stimuli used at test. In previous studies, an LFE on talker discrimination (not talker learning or recognition) has been reported when listeners were tested on monosyllabic words presented in isolation (Levi, 2017; Levi and Schwartz, 2013; Winters *et al.*, 2008) but not when listeners were tested on sentences (Wester, 2012, and the current study). However, the lack of an LFE could also be explained by differences in the tasks used in the present study and earlier studies of the LFE. Compared to the same-different task used in Experiment 1, which merely requires listeners to compare two utterances presented in quick succession, the tasks that are commonly used in the LFE literature (talker learning and voice line-up tasks) differ substantially with regard to task demands and memory requirements. To test for the possibility that task differences account for the absence of an LFE in Experiment 1, we tested a new group of adults and children (drawn from the same participant pool) on a different paradigm in Experiment 2 while keeping other important factors (like talkers, sentences, visual stimuli, and experimental setup) constant across experiments. This approach allowed us to test for potential task effects on the LFE while revisiting our original research question: Does the strength of the LFE increase with age?

III. EXPERIMENT 2

In Experiment 2, we used a child-friendly variant of a "voice line-up" talker recognition task (rather than an AX "same-different" talker discrimination task) to test adults' and children's ability to identify English–Polish bilinguals speaking either a familiar language (English) or an unfamiliar language (Polish). We expected adults to outperform children and both age groups to exhibit an LFE (i.e., to perform better for English than Polish). Consistent with our main research question, we moreover explored whether the LFE would be stronger in children than adults.

A. Methods

1. Participants

Thirty-two native English-speaking adults ($M_{age} = 19.1$ years, SD = 2.1; 24 female) and 32 5- and 6-year-old children ($M_{age} = 6.0$ years, SD = 0.4; 16 female) from the Greater Toronto Area were tested. Eligibility criteria were the same as in Experiment 1. Data for two additional children were excluded due to children's unwillingness to complete the study as instructed.

2. Stimuli

Auditory stimuli consisted of a subset of the recorded sentences from Experiment 1. Eight of the alien characters used in Experiment 1 were used as visual stimuli. One character was assigned per voice per language condition.

3. Procedure

Participants were tested individually. The software used to design and administer the experiment, and the experimental setup used for testing, were the same as in Experiment 1.

In Experiment 2, instead of determining whether two sentences were produced by the same talker or two different talkers, participants were exposed to a talker speaking either English or Polish, and after a short delay, were asked to identify which of two talkers presented to them in a voice line-up matched the talker they had heard before. Training and test sentences were always spoken in the same language. Sentences were never repeated within and across trials, requiring listeners to generalize talker-specific information across sentences.

More specifically, each trial of the experiment consisted of three stages (see Fig. 3). In Stage 1 (familiarization), an alien appeared on the screen and participants heard four sentences spoken by the same talker [two different sentences, each repeated twice; interstimulus interval (ISI) between sentences = 300 ms; ISI between the repetition of the set = 500 ms]. In Stage 2 (retention), participants watched a 1 min long movie showing episodes extracted from popular children's TV shows and movies (dubbed with music and sound effects but no speech), which was intended to introduce a delay and remove (at least partly) auditory memory traces. In Stage 3 (recall), two spaceships appeared on the screen. When participants touched one of the spaceships, they heard a sentence spoken by one of the four talkers, and when they touched the other spaceship, they heard a different sentence spoken by another talker. Children were told that the alien they saw earlier was hiding in one of the spaceships. To find out where the alien was hiding, children had to decide which of the two voices in the spaceships sounded like the alien they saw in Stage 1. Participants could listen as often as they wished to the voices in Stage 3, and were advised to listen to both voices before making a decision.

The target voice was always present (closed-set line-up). Participants made their choice by dragging the appropriate spaceship onto the image of a planet presented at the top of the screen. They received feedback on the correctness of their responses (a smiling face was shown for every correct response and a frowning face for every incorrect response). The last two trials always showed a smiling face (irrespective of performance) to ensure that the children left the lab feeling good about their performance.

Participants completed one practice trial in English (consisting of acoustically-distinct voices that were not included in the experimental trials) before completing eight experimental trials (with each trial lasting around 2 min). After four trials, participants took a timed break for 4 min during which adults played Sudoku and children played a (not speech- or language-related) game of their choice with the experimenter (e.g., Snakes and Ladders, Tic-tac-toe). The experiment was self-paced; participants took 20–25 min to complete the experiment.

The order of the eight experimental trials (four English, four Polish) was counterbalanced across participants. In each language condition, each voice was presented once as the target and once as the distractor. The side of presentation of the target voice during test (i.e., whether the target was linked to the left or right spaceship) was randomized.

B. Results and discussion

We analyzed mean proportion correct talker recognition (see Fig. 4). Both adults ($M_{\text{English}} = 0.84$, SD = 0.19) and children ($M_{\text{English}} = 0.65$, SD = 0.22) performed significantly above chance level (0.5) for English, and adults ($M_{\text{Polish}} = 0.68$, SD = 0.24) also did so for Polish (all $p < 0.001$); children ($M_{\text{Polish}} = 0.54$, SD = 0.21) performed at chance for Polish ($p = 0.305$).

To assess adults' and children's talker recognition performance, we used a generalized mixed-effects model with the same fixed and random effects structure as specified for Experiment 1. The model revealed a significant main effect of age ($\beta = 0.82$, SE = 0.21, $z = 3.99$, $p < 0.001$), with adults performing better than children, and a significant main effect of language ($\beta = 0.68$, SE = 0.20, $z = 3.34$, $p < 0.001$), with adults and children alike recognizing the talkers more accurately when the talkers spoke English than when they spoke Polish. The age × language interaction was not significant ($p = 0.289$).

As in Experiment 1, adults performed better than children. But at the same time, children's performance (in particular for English) was again better than expected, especially when considering that the talkers were unfamiliar to the children and did not differ in age or gender (and these factors have been found to hinder talker learning in young children; Creel and Jiménez, 2012).

In line with our predictions, and contrary to our findings in Experiment 1, both adults and children performed better when the talkers spoke the familiar language than when they spoke the unfamiliar language, thus providing clear evidence of an LFE on talker recognition. Since important factors like talkers, stimuli and listener ages were kept consistent across experiments, task differences between Experiment 1 (talker discrimination) and Experiment 2 (talker recognition) most likely account for the fact that we observed an LFE in the second but not the first experiment. The LFE was absent in Experiment 1, which only involved an acoustic-phonetic comparison of two consecutively-presented voices, but a robust LFE emerged in Experiment 2, where listeners were required to hold talker-specific detail in memory for considerably longer.

Although we found a robust LFE in Experiment 2, the size of the effect did not differ between adults and children (no significant age × language interaction). However, descriptively speaking, the LFE was slightly stronger for adults than children (see Fig. 4). This led us to hypothesize that the LFE might be stronger for the older children than the younger children in our 5- and 6-year-old test population. And indeed, when we split the 2-year age range in half and compared LFE strength in children above age 6 (M_{age} = 6.3 years, SD = 0.2; 7 female) and below age 6 (M_{age} = 5.7 years, SD = 0.3; 9 female), we found that the LFE for the children was in fact driven by the older children. The 6-year-olds showed better talker recognition for English ($M = 0.66$, SD = 0.20) than Polish ($M = 0.48$, SD = 0.23), $t(15) = 2.30$, $p = 0.036$, but the 5-year-olds performed equally well for English ($M = 0.64$, SD = 0.24) and Polish ($M = 0.59$, SD = 0.18), $t(15) = 0.72$, $p = 0.485$ (see Fig. 5). Would this pattern of results replicate? Do 5-year-olds and 6-year-olds truly differ so strongly in their ability to recognize native- and foreign-language speakers? In Experiment 3, we tested a new group of children on the same task as in Experiment 2 to determine whether the observed age difference can be replicated (note that this approach also addresses recent calls in psychology to more frequently replicate empirical findings; e.g., Open Science Collaboration, 2015).

IV. EXPERIMENT 3

In Experiment 2, both children and adults showed an LFE, but *post hoc* comparisons revealed that the effect in the

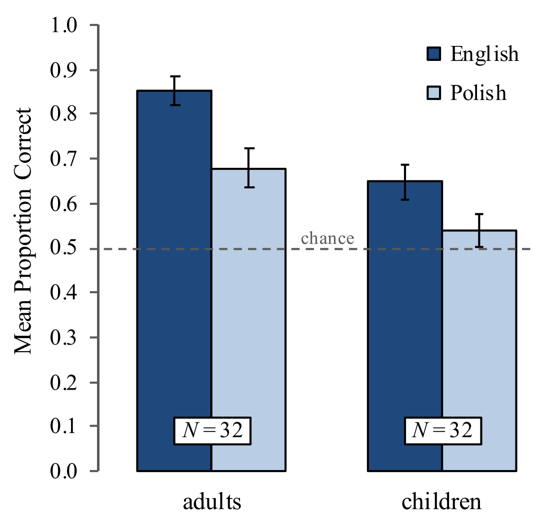

FIG. 4. (Color online) Mean proportion correct for English and Polish in Experiment 2. Both age groups exhibited a robust LFE on talker recognition (error bars indicate SE).

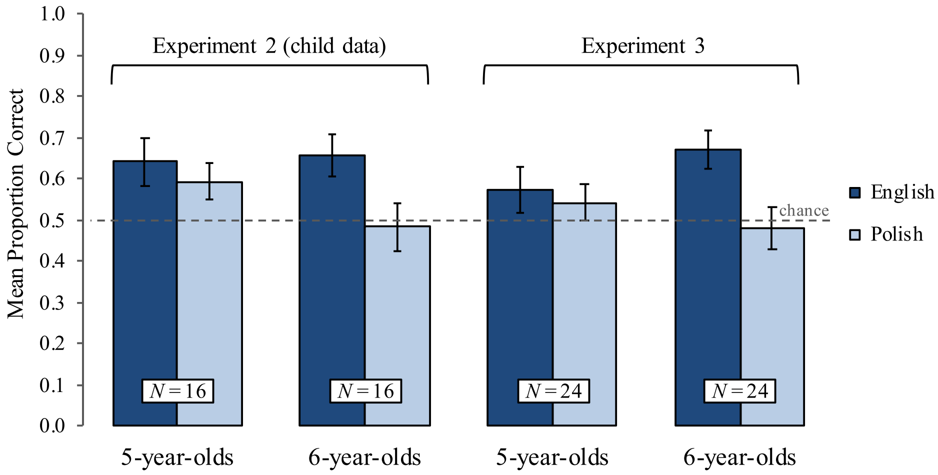

FIG. 5. (Color online) Mean proportion correct for English and Polish in Experiment 2 (child data only) and Experiment 3. The 6-year-olds, but not the 5-year-olds, showed a robust LFE on talker recognition (error bars indicate SE).

children was primarily driven by the 6-year-olds (i.e., the 5-year-olds performed equally well for the native and foreign language). This finding, albeit unexpected, was intriguing because it alludes to important developmental differences between younger and older children and thus potentially points to an explanation for how mature talker recognition abilities develop. The goal of Experiment 3 was to establish whether the age effect in Experiment 2 can be replicated in a different group of 5- and 6-year-old children.

A. Method

1. Participants

Twenty-four 5-year-old ($M_{\text{age}} = 5.5$ years, SD = 0.2; 11 female) and 24 6-year-old ($M_{\text{age}} = 6.5$ years, SD = 0.3; 12 female) monolingual English-speaking children from the Greater Toronto Area were tested. Eligibility criteria were the same as in Experiments 1 and 2. Seven additional children were tested, but their data were excluded due to their unwillingness to complete the study as instructed (6) or technical error (1).

2. Stimuli

Stimuli were the same as the stimuli used in Experiment 2.

3. Procedure

The procedure was identical to the procedure in Experiment 2.

B. Results and discussion

We calculated mean proportion correct talker recognition in English and Polish for the 5- and 6-year-olds separately (only here the comparison between these two age groups was planned rather than *post hoc*). As in Experiment 2, the 6-year-olds were significantly better at recognizing talkers speaking English ($M = 0.67$, SD = 0.24) than Polish ($M = 0.48$, SD = 0.23), $t(23) = 3.09$, $p = 0.005$, but the performance of the 5-year-olds did not significantly differ between English ($M = 0.57$, SD = 0.27) and Polish ($M = 0.54$, SD = 0.22), $t(23) = 0.41$, $p = 0.689$.

We then collapsed the data from both experiments and assessed children's talker recognition performance using a generalized mixed-effects model with age (0.5: 6-year-olds, −0.5: 5-year-olds), language (0.5: English, −0.5: Polish) and the age × language interaction as fixed effects, as well as random intercepts for participant and talker, and a random slope for language by participant. This model revealed a significant main effect of language ($\beta = 0.48$, SE = 0.16, $z = 2.96$, $p = 0.003$) but no main effect of age ($p = 0.867$). In line with our goal (to determine whether LFE strength differs between 5- and 6-year-olds), the interaction between age and language was significant ($\beta = 0.65$, SE = 0.32, $z = 2.00$, $p = 0.045$). This result suggests that the language spoken at test indeed affected talker recognition in the older and younger children differently. To follow up on this finding, we constructed separate models for the two age groups, with language as fixed effect, random intercepts for participant and talker, and a by-participant random slope for language. We found a significant effect of language for the 6-year-olds ($\beta = 0.80$, SE = 0.23, $z = 3.48$, $p < 0.001$), with children performing better for English than Polish. However, there was no effect of language for the 5-year-olds ($p = 0.483$), suggesting that these children performed equally well for English and Polish.

To summarize, the combined data from Experiments 2 and 3 showed that the older children tested in our study exhibited a strong native-language benefit for talker recognition, but the younger children showed no evidence of an LFE. The difference in LFE strength between children at age 5 and age 6 suggests that the LFE grows stronger between infancy and childhood, and that children might reach adult-like levels of the LFE earlier in childhood than we predicted at the outset of the study. We discuss the implications of these findings for our understanding of the development of mature talker recognition abilities in Sec. V.

V. GENERAL DISCUSSION

Although many studies have shown that both talker recognition and spoken language processing skills continue to improve into late childhood, very few studies have considered that the development of these two abilities might be related. Here, we examined the LFE—a hallmark of mature talker recognition—in both children and adults. We reasoned that if the development of talker recognition and the development of spoken language processing are related, then the LFE should be weaker in young children than adults because children do not yet have adult-like language competence. In all three reported experiments, we replicated past studies by showing superior talker recognition in adults compared to children. In addition, we found that only adults and 6-year-olds, but not 5-year-olds, recognized talkers who speak a familiar language better than talkers who speak an unfamiliar language (i.e., all but the youngest age group showed an LFE). We conclude that the LFE appears to grow stronger over the course of early childhood development; as children's spoken language processing skills improve, their talker recognition abilities also become more adult-like.

While the current study provides evidence that the LFE strengthens with age, the developmental pattern that we observed is not what we initially expected. At the outset of this study, we hypothesized that 5- and 6-year-old children on the whole might show a weaker LFE than adults, but this is not what we found. Instead, we found an increase in LFE strength between 5 and 6 years of age (with only the older children showing an LFE). This suggests that the attainment of a robust LFE might occur earlier than we originally predicted. However, does this finding necessarily imply that the LFE reaches adult-like levels by 6 years of age? Although the size of the LFE in adults and 6-year-olds tested in our study did not differ,[2] we cannot for certain conclude that the LFE is mature by age 6. Indeed, since language competence continues to improve past 6 years of age, it seems possible that given a different (potentially more sensitive) experimental paradigm we might find further strengthening of the LFE between age 6 and adulthood. But setting aside the issue of whether the size of the LFE continues to increase past age 6, the question of why LFE strength differed so strongly between the 5- and 6-year-olds in our study remains. How can children who are, on average, just one year apart in age show such a dramatic difference in the way they process talker identity information in speech?

The difference in LFE strength between the younger and older children could be related to a wide range of non-linguistic factors that may be correlated with age, such as general cognitive maturation or auditory-perceptual processing abilities. However, we noticed upon reflecting on the specific age group that we tested that the difference in talker recognition in 5- and 6-year-olds could also be linked to literacy education in schools and at home. During pre school and early school years, children acquire and continuously improve their reading and writing skills. In doing so, children learn that words are composed of smaller meaningful units (phonemes) that can signal differences in meaning (like the onset in *bear* and *wear*; e.g., Snowling and Hulme, 1994). Bregman and Creel (2014) previously suggested that as children's representations of native-language phonemes mature, they will begin to encode talker variation more proficiently. Here, we refine this notion by proposing that the heightened awareness of the segmental structure of language might additionally cause a shift in attention to different types of speech-related information to recognize talkers. As proposed by Levi and Schwartz (2013), listeners might weigh language-dependent and language-independent information in speech differently across development (see also Winters *et al.*, 2008). As children grow older, they might rely more heavily on language-dependent information (e.g., a talker's idiolect, dialect or sociolect) at the expense of language-independent information (like pitch or voice quality).

In the current study, we established that the LFE appears to strengthen in early school years, but what factors are driving these differences across development is currently less clear. Does the LFE emerge because children get better at recognizing speakers of their native language or because they get worse at recognizing speakers of a foreign language (or both)? This study was not explicitly designed to address this question, but patterns in our data inspire us to speculate on what the most likely answer might be. In both Levi and Schwartz (2013) and in the current study, we see that children get better at native-language talker recognition and worse at foreign-language talker recognition, but in neither study does this pattern reach statistical significance. Future work should be designed explicitly to examine whether this pattern is a coincidence or truly reflects a developmental difference between younger and older children. Indeed, ongoing work in our lab has once again shown this pattern in 5- versus 6-year-old children tested on an entirely different set of stimuli. Following up this work is important because if improvements in native-language talker recognition are in fact accompanied by a decline in foreign-language talker recognition across development, then this will provide a clue to the underlying mechanisms driving the LFE and the development of mature talker recognition skills more generally. And interestingly, it would also suggest that the development of expertise in talker recognition in some ways parallels the development of expertise in speech perception (where improvements in native-language speech perception come at the cost of speech perception in non-native languages; e.g., Werker and Tees, 1984; Kuhl *et al.*, 2003).

Finally, the present study suggests that the strength of the LFE not only depends on the listener's age but also on the task demands and memory requirements of the perceptual tasks used at test (see also Creel and Jiménez, 2012; Levi, 2017; Winters *et al.*, 2008). In Experiment 1, listeners of all age groups successfully discriminated between the talkers irrespective of the language spoken (no LFE). Listeners could perform this task simply by comparing two utterances on an acoustic-phonetic level without necessarily requiring access to linguistic information. However, when we changed the paradigm from discrimination in Experiment 1 to recognition in Experiments 2 and 3 (while keeping other factors constant), the language spoken by the talkers determined listeners' success or failure in identifying the talkers (robust LFE). In contrast to Experiment 1, the recognition task involved a greater load on auditory memory and listeners were required to store talker-related information for much longer (note that similar task effects have been reported for other domains of speech processing, such as phoneme perception; e.g., Gerrits and Schouten, 2004; Sadakata and McQueen, 2013). Based on these findings, we cautiously propose that the locus of the native-language benefit for talker recognition is in the later (encoding, storage, or retrieval) stages in perception where linguistic knowledge is applied. Future research will be needed to further clarify the role of listener age (and by inference language experience) and task demands in the emergence of the LFE, and to investigate whether the observed age and task effects interact across development (see, e.g., Creel and Quam, 2015).

VI. CONCLUSIONS

The current study examined whether the strength of the LFE differs between adults and children. Results showed that the LFE appears to grow stronger in early childhood, suggesting that children at different ages attend to different types of linguistic information to determine a talker's identity. In addition, this study found that the emergence of the LFE seems to be task-dependent. Adults and older children elicited a robust LFE in a talker recognition task (which involved a considerable memory component), but neither age group showed an LFE in a talker discrimination task. This work helps us to better understand the development of mature talker recognition and spoken language processing skills in young children, and how these two domains interact across early childhood development. The current findings motivate future studies on the development of the LFE as a means to further advance our understanding of the relationship between language learning and talker learning across language and ages, and thus of human speech processing more generally.

ACKNOWLEDGMENTS

The authors thank all the speakers for producing the speech stimuli, Katrina Aranas for assistance with stimulus creation and data collection, Jessamyn Schertz and Angela Cooper for statistical advice, and all participating families. This research was supported by grants awarded to E.K.J. 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: EXAMPLE 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.)

[1]Results did not differ when we included the ten children that had been excluded due to a response bias ($M_{\text{English}} = 0.61$, SD = 0.15; $M_{\text{Polish}} = 0.63$, SD = 0.19). Children still performed significantly above chance for both languages, adults still outperformed children, and children still failed to show an LFE.

[2]Statistical comparison of the performance of the adults ($M_{\text{English}} = 0.84$, SD = 0.19; $M_{\text{Polish}} = 0.68$, SD = 0.24) and 6-year-olds ($M_{\text{English}} = 0.67$, SD = 0.21; $M_{\text{Polish}} = 0.48$, SD = 0.24) showed no difference in LFE strength between these two age groups (no significant age × language interaction; $p = 0.861$).

Andics, A., McQueen, J. M., and Van Turennout, M. (**2007**). "Phonetic context influences voice discriminability," in *Proceedings of the 16th International Congress of Phonetic Sciences*, August 6–10, Saarbrücken, Germany, pp. 1829–1832.
Bartholomeus, B. (**1973**). "Voice identification by nursery school children," Can. J. Psychol. **27**(4), 464–472.
Bates, D., Mächler, M., Bolker, B., and Walker, S. (**2015**). "Fitting linear mixed-effects models using lme4," J. Stat. Softw. **67**(1), 1–48.
Bregman, M. R., and Creel, S. C. (**2014**). "Gradient language dominance affects talker learning," Cognition **130**(1), 85–95.
Creel, S. C., and Jiménez, S. R. (**2012**). "Differences in talker recognition by preschoolers and adults," J. Exp. Child Psychol. **113**(4), 487–509.
Creel, S. C., and Quam, C. (**2015**). "Apples and oranges: Developmental discontinuities in spoken-language processing?," Trends Cogn. Sci. **19**(12), 713–716.
Cutler, A., Andics, A., and Fang, Z. (**2011**). "Inter-dependent categorization of voices and segments," in *Proceedings of the 17th International Congress of Phonetic Sciences*, August 17–21, Hong Kong, China, pp. 552–555.
Fecher, N., and Johnson, E. K. (**2018**). "The native-language benefit for talker identification is robust in 7.5-month-old infants," J. Exp. Psychol. Learn. (in press).
Fleming, D., Giordano, B. L., Caldara, R., and Belin, P. (**2014**). "A language-familiarity effect for speaker discrimination without comprehension," Proc. Natl. Acad. Sci. USA **111**(38), 13795–13798.
Gerrits, E., and Schouten, M. E. H. (**2004**). "Categorical perception depends on the discrimination task," Percept. Psychophys. **66**(3), 363–376.
Goggin, J. P., Thompson, C. P., Strube, G., and Simental, L. R. (**1991**). "The role of language familiarity in voice identification," Mem. Cogn. **19**(5), 448–458.
Goldinger, S. D. (**1996**). "Words and voices: Episodic traces in spoken word identification and recognition memory," J. Exp. Psychol. Learn. **22**(5), 1166–1183.
Hazan, V., and Barrett, S. (**2000**). "The development of phonemic categorization in children aged 6-12," J. Phon. **28**(4), 377–396.
Johnson, E. K., Bruggeman, L., and Cutler, A. (**2018**). "Abstraction and the (misnamed) Language Familiarity Effect," Cogn. Sci. **42**(2), 633–645.
Johnson, E. K., Westrek, E., Nazzi, T., and Cutler, A. (**2011**). "Infant ability to tell voices apart rests on language experience," Dev. Sci. **14**(5), 1002–1011.
Johnson, K. (**1990**). "The role of perceived speaker identity in F0 normalization of vowels," J. Acoust. Soc. Am. **88**(2), 642–654.
Kadam, M. A., Orena, A. J., Theodore, R. M., and Polka, L. (**2016**). "Reading ability influences native and non-native voice recognition, even for unimpaired readers," J. Acoust. Soc. Am. **139**(1), EL6–EL12.
Köster, O., and Schiller, N. O. (**1997**). "Different influences of the native language of a listener on speaker recognition," Int. J. Speech Lang. Law **4**(1), 18–28.
Kuhl, P. K., Tsao, F.-M., and Liu, H.-M. (**2003**). "Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning," Proc. Natl. Acad. Sci. USA. **100**(15), 9096–9101.
Levi, S. V. (**2017**). "Another bilingual advantage? Perception of talker-voice information," Biling. Lang. Cogn. (published online 2017).
Levi, S. V., and Schwartz, R. G. (**2013**). "The development of language-specific and language-independent talker processing," J. Speech Lang. Hear Res. **56**(3), 913–920.
Mann, V. A., Diamond, R., and Carey, S. (**1979**). "Development of voice recognition: Parallels with face recognition," J. Exp. Child Psychol. **27**(1), 153–165.
Mullennix, J. W., Pisoni, D. B., and Martin, C. S. (**1989**). "Some effects of talker variability on spoken word recognition," J. Acoust. Soc. Am. **85**(1), 365–378.

Nittrouer, S. (**1992**). "Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries," J. Phon. **20**(3), 351–382.
Nygaard, L. C., and Pisoni, D. B. (**1998**). "Talker-specific learning in speech perception," Percept. Psychophys. **60**(3), 355–376.
Ohde, R. N., and Haley, K. L. (**1997**). "Stop-consonant and vowel perception in 3- and 4-year-old children," J. Acoust. Soc. Am. **102**(6), 3711–3722.
Open Science Collaboration. (**2015**). "Estimating the reproducibility of psychological science," Science **349**(6251), 4716.
Orena, A. J., Theodore, R. M., and Polka, L. (**2015**). "Language exposure facilitates talker learning prior to language comprehension, even in adults," Cognition **143**, 36–40.
Perea, M., Jiménez, M., Suárez-Coalla, P., Fernández, N., Viña, C., and Cuetos, F. (**2014**). "Ability for voice recognition is a marker for dyslexia in children," Exp. Psychol. **61**(6), 480–487.
Perrachione, T. K. (**2017**). "Speaker recognition across languages," in *The Oxford Handbook of Voice Perception*, edited by S. Frühholz and P. Belin (Oxford University Press, Oxford, United Kingdom).
Perrachione, T. K., Del Tufo, S. N., and Gabrieli, J. D. E. (**2011**). "Human voice recognition depends on language ability," Science **333**(6042), 595.
Perrachione, T. K., Dougherty, S. C., McLaughlin, D. E., and Lember, R. A. (**2015**). "The effects of speech perception and speech comprehension on talker identification," in *Proceedings of the 18th International Congress of Phonetic Sciences*, August 10–14, Glasgow, United Kingdom.
R Core Team (**2016**). "R: A language and environment for statistical computing [computer program]," https://www.r-project.org/ (Last viewed October 25, 2017).
Sadakata, M., and McQueen, J. M. (**2013**). "High stimulus variability in nonnative speech learning supports formation of abstract categories: Evidence from Japanese geminates," J. Acoust. Soc. Am. **134**(2), 1324–1335.
Snowling, M., and Hulme, C. (**1994**). "The development of phonological skills," Philos. Trans. R. Soc. Lond. B **346**(1315), 21–27.
Spence, M. J., Rollins, P. R., and Jerger, S. (**2002**). "Children's recognition of cartoon voices," J. Speech Lang. Hear. Res. **45**(1), 214–222.
Sullivan, K. P. H., and Schlichting, F. (**2000**). "Speaker discrimination in a foreign language: First language environment, second language learners," Int. J. Speech Lang. LA **7**(1), 95–111.
Thompson, C. P. (**1987**). "A Language Effect in Voice Identification," Appl. Cogn. Psych. **1**(2), 121–131.
Van Heugten, M., Volkova, A., Trehub, S. E., and Schellenberg, E. G. (**2014**). "Children's recognition of spectrally degraded cartoon voices," Ear Hear. **35**(1), 118–125.
Werker, J. F., and Tees, R. C. (**1984**). "Cross-language speech perception: Evidence for perceptual reorganization during the first year of life," Infant. Behav. Dev. **7**(1), 49–63.
Wester, M. (**2012**). "Talker discrimination across languages," Speech Commun. **54**(6), 781–790.
Winters, S. J., Levi, S. V., and Pisoni, D. B. (**2008**). "Identification and discrimination of bilingual talkers across languages," J. Acoust. Soc. Am. **123**(6), 4524–4538.