



Brief article

Language exposure facilitates talker learning prior to language comprehension, even in adults

Adriel John Orena^{a,b,*}, Rachel M. Theodore^c, Linda Polka^{a,b}^a School of Communication Sciences and Disorders, McGill University, Canada^b Centre for Research on Brain, Language and Music, McGill University, Canada^c Department of Speech, Language, and Hearing Sciences, University of Connecticut, United States

ARTICLE INFO

Article history:

Received 16 January 2015

Revised 18 May 2015

Accepted 1 June 2015

Available online 23 June 2015

Keywords:

Speech perception

Talker identification

Talker learning

Language exposure

ABSTRACT

Adults show a native language advantage for talker identification, which has been interpreted as evidence that phonological knowledge mediates talker learning. However, infants also show a native language benefit for talker discrimination, suggesting that sensitivity to linguistic structure due to systematic language exposure promotes talker learning, even in the absence of functional phonological knowledge or language comprehension. We tested this hypothesis by comparing two groups of English-monolingual adults on their ability to learn English and French voices. One group resided in Montréal with regular exposure to spoken French; the other resided in Storrs, Connecticut and did not have French exposure. Montréal residents showed faster learning and better retention for the French voices compared to their Storrs-residing peers. These findings demonstrate that systematic exposure to a foreign language bolsters talker learning in that language, expanding the gradient effect of language experience on talker learning to perceptual learning that precedes sentence comprehension.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The speech signal conveys both the talker's communicative message and identity, with a growing body of evidence indicating that these two components are fundamentally intertwined (Creel & Bregman, 2011). Listeners receive language comprehension benefits for familiar compared to unfamiliar talkers (e.g., Clarke & Garrett, 2004; Nygaard, Sommers, & Pisoni, 1994; Theodore & Miller, 2010), demonstrating that experience with a talker's voice facilitates recovery of the communicative message. Listeners also show a talker learning advantage for native compared to nonnative talkers (e.g., Goggin, Thompson, Strube, & Simental, 1991; Johnson, Westrek, Nazzi, & Cutler, 2011; Winters, Levi, & Pisoni, 2008), suggesting that access to the communicative message likewise facilitates talker identification. The goal of the current work is to contribute to a theoretical understanding of how language experience influences talker identification.

The literature confirms a strong role for phonological processing in mediating talker identification. Phonological processing refers to knowledge of the abstract sound structure of a specific language.

Studies of adult second-language learners suggest that a gradient improvement in phonological knowledge of a language can improve talker identification in that language (Köster & Schiller, 1997; Sullivan & Schlichting, 2000). In addition, Bregman and Creel (2014) found that early learners of a second language were more skilled than late learners with respect to talker learning in their second language. Converging evidence that phonological competence influences talker identification comes from studies of adults with dyslexia, who demonstrate poor talker identification even in their native language, with performance related to their degree of phonological deficit (Perrachione, Del Tufo, & Gabrieli, 2011).

However, developmental research suggests that functional phonological knowledge or language comprehension are not the sole aspects of language experience that influence talker identification. Newborns can recognize their mother's voice over another female's voice (e.g., DeCasper & Fifer, 1980). Moreover, 7-month-old infants who have yet to develop a mature phonological system already show some native language benefit for talker discrimination (Johnson et al., 2011). These findings suggest that infants' sensitivity to the structure of their native language – which precedes phonological competence and word comprehension – may promote talker learning. This leads to the hypothesis that adults with regular exposure to a nonnative language may also promote a native-like benefit for talker learning in that nonnative language. Indeed, one notable exception to the native language

* Corresponding author at: School of Communication Sciences and Disorders, McGill University, 2001 McGill College, 8th Floor, Montréal, Québec H3A 1G1, Canada.

E-mail address: adriel.orena@mail.mcgill.ca (A.J. Orena).

benefit for talker identification was reported by Johnson et al. (2011). They found that English-monolingual adults living in Toronto performed equally well at identifying voices in English and in Italian and hypothesized that this could be an effect of language exposure given the large Italian community in Toronto.

Here we examined whether phonological competence in a language is necessary for talker learning or, alternatively, whether having regular exposure to talkers of a novel language is sufficient. In contrast to previous studies that have shown a language familiarity effect in individuals after they have gained some expressive and receptive proficiency in the language (e.g., Köster & Schiller, 1997; Sullivan & Schlichting, 2000), here we asked whether having exposure to an unfamiliar language without gaining proficiency (à la monolinguals living in bilingual communities) could promote talker learning in that language. If exposure to nonnative acoustic–phonetic variation promotes talker-learning benefits, then listeners who overhear an unfamiliar language will show increased talker identification compared to listeners who do not overhear that language. If, however, more sophisticated functional language knowledge is needed, then both types of listeners should show similar performance on the talker-learning task.

2. Methods

2.1. Participants

Two groups of English-monolinguals were recruited from cities with different language environments: Montréal, Québec and Storrs, Connecticut. The latest census (2011) reports that 55.8% of individuals living in Montréal have some fluency in both English and French. The two languages co-exist in many cultural and social aspects in Montréal. Thus, Montréal residents have experience hearing French spoken by different talkers. In stark contrast, the latest census (2000) for Storrs, Connecticut reports that 1.3% of the population has any degree of fluency in French. Storrs is an overwhelmingly English-monolingual environment where residents have virtually no opportunity to hear French. Sampling English monolinguals residing in these two communities provides a way to control fluency in English while manipulating exposure to French. Responses to the Language Experience and Proficiency Questionnaire (Marian, Blumenfeld, & Kaushanskaya, 2007) and a questionnaire developed in our laboratory confirmed that exposure to French was as predicted by residence.

Participants from Montréal included 20 adults (18–32 years; 7 males). Participants from Storrs included 19 adults (18–27 years; 8 males). One additional participant from Storrs was excluded from analyses due to fluency in French. All acquired English from birth and lived in North American English-monolingual communities prior to residing in Montréal or Storrs. Participants rated their exposure to French on a 5-point scale (1 = never, 5 = daily). All Montréal participants reported some level of exposure to spoken French (mean = 2.65, SD = 1.18). Only 10 out of the 19 Storrs participants reported any French exposure; the French exposure rating among this subset (mean = 1.60, SD = 0.70) was significantly lower than the Montréal group [$t(28) = 2.58$, $p = .015$, $d = 0.97$]. All 20 participants in the Montréal group and 8 out of 19 in the Storrs group had received formal classroom instruction in French. Nonetheless, both groups could neither understand spoken French or converse in French, and none were able to understand the sentences presented in the experiment.

2.2. Stimuli

The auditory stimuli are described in Valji (2004). In brief, four female native speakers of Canadian-English and four female native

speakers of Canadian-French were recorded producing 10 sentences in their native language. For each language, five of the sentences were used during training and test, and the remaining five were used only during test. Acoustic analyses were used to calculate mean duration, fundamental frequency (F_0), and F_0 standard deviation for each speaker. We followed methodology outlined in Johnson et al. (2011) to calculate a ratio of the variances in each of these dimensions among the English and French speakers. The within-language variability was not significantly different between the two languages for all three dimensions [$(F(3,3) < 9.1$ for $\alpha = .05$ in all cases], indicating that the speakers in each language should be equally difficult to distinguish based on these acoustic parameters.

Pilot tests confirmed that the English and French stimuli were equally “easy” to learn for native speakers of each language. We tested 10 native English speakers and 10 native French speakers (some were French–English bilinguals, but all were French dominant and did not learn English before the age of 10) on a talker-learning task in their native language (identical to the training phase described below). There was no difference in performance between the two groups in terms of number of blocks to reach criterion [English: $M = 3.30$, $SE = .60$; French: $M = 3.80$, $SE = .61$; $t(18) = .59$, $p > .250$; $d = .28$]. Since baseline performance for our stimuli in native-language conditions was equal across the two languages, any difference in performance across English and French language conditions can be attributed to language exposure/experience rather than stimulus properties.

2.3. Procedure

The design of this experiment was adapted from Bregman and Creel (2014). Participants were tested individually in a single, 2-h session in a sound-attenuated space. Visual stimuli were presented on a computer monitor and auditory stimuli were presented over headphones (Sony MDR-V6) at a constant, comfortable listening level. All responses were collected via button box. Testing procedures and conditions were identical across both laboratories.

Participants learned to identify four English talkers and four French talkers. These voices were associated with one of four cartoon avatars (Fig. 1). For each language, participants completed a training phase followed by a test phase (Fig. 2). Each language was tested separately, with language order counter-balanced within each location group.

The training phase provided listeners with an opportunity to learn the voice-avatar pairings. Each block consisted of 60 randomized trials (4 talkers \times 5 sentences \times 3 repetitions). On each trial, auditory and visual stimuli were simultaneously presented. The auditory stimulus consisted of one of the sentences. The visual stimulus was an array of the four cartoon faces in a single row, with arrangement held constant throughout the experiment. Participants were asked to indicate the identity of the voice by pressing one of four labeled buttons. Feedback was provided on each trial, with the word “Correct” or “Incorrect” appearing on the screen with the correct avatar. The next trial began after 2000 ms pause. The training block was repeated until the participant met learning criterion, defined as 85% correct talker identification within a single block (52/60 trials) or the completion of eight training blocks. The dependent measure was the number of the training blocks required to meet the learning criterion.

Following the training phase, participants were tested on their ability to retain the voice and face pairings presenting during the training phase. The test included the sentences presented during training along with novel sentences produced by the same talkers in order to assess generalization. The test phase consisted of 120 randomized trials (4 talkers \times 10 sentences (5 trained and 5 novel) \times 3 repetitions). As schematized in Fig. 2, procedure for the test

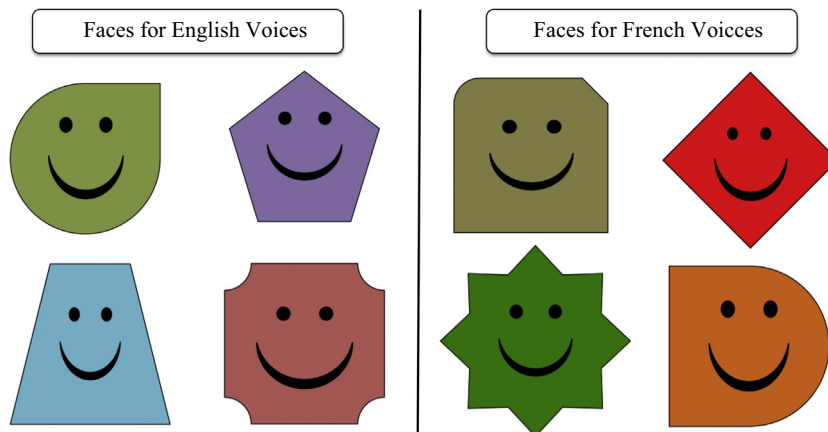


Fig. 1. Cartoon faces used for voice association with the English and French speakers.

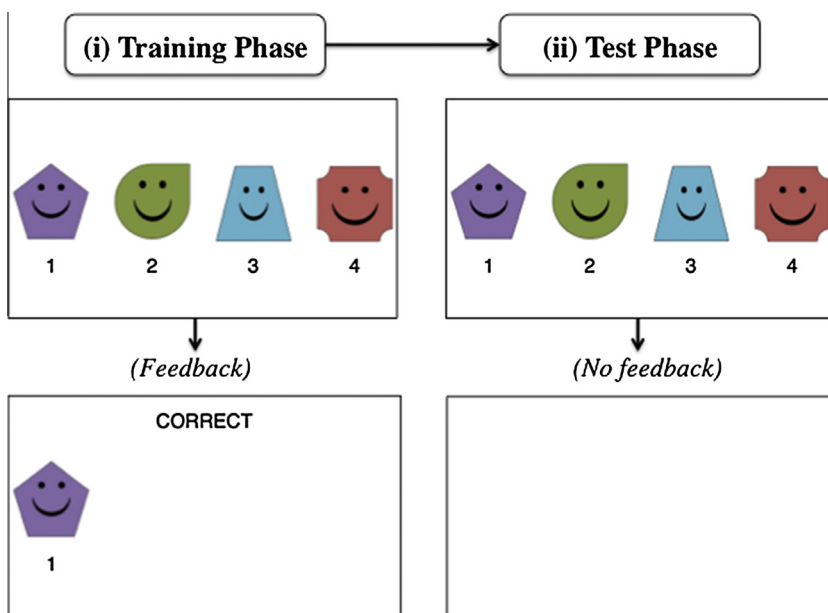


Fig. 2. Schematic of the training and test phases.

phase was identical to the training phase except that no feedback was provided. The dependent measure was percent correct talker identification.

3. Results

Learning rate for each language was measured as the number of blocks required to reach the learning criterion, which is shown separately for participants in Montréal and Storrs in Fig. 3. Mean number of training blocks was submitted to ANOVA with the within-subjects factor of stimulus language (English versus French) and the between-subjects factor of location (Montréal versus Storrs). As expected, there was a main effect of stimulus language with fewer training blocks required to reach criterion for English compared to French [$F(1,37) = 101.09, p < .001, \eta_p^2 = 0.73$]. There was also a main effect of location, with fewer training blocks required for those in Montréal compared to those in Storrs [$F(1,37) = 4.36, p = .044, \eta_p^2 = 0.11$]. Critically, there was an interaction between stimulus language and location [$F(1,37) = 9.41, p = .004, \eta_p^2 = 0.20$]. Planned comparisons showed that there was no difference in the number of training blocks between the two

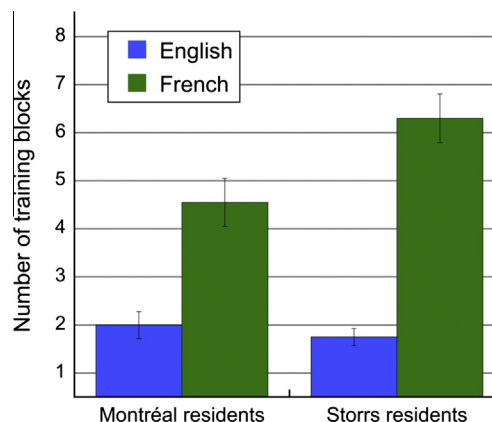


Fig. 3. Mean number of training blocks to reach performance criterion in English and French for residents of Montréal and Storrs. Error bars indicate standard error of the mean.

locations for English [$t(37) = 0.77, p > .250, d = 0.25$], but that there was a difference for French, with participants in Montréal requiring significantly fewer blocks to reach criterion compared to

participants in Storrs [$t(37) = -2.83$, $p = .007$, $d = -0.93$]. Thus, English monolinguals who had frequent exposure to French learned to identify novel French talkers faster than listeners with little or no exposure to French.

Test performance (Fig. 4) was measured in terms of percent correct talker identification and was calculated separately for the two languages and the two types of test stimuli (trained and novel items). Mean percent correct talker identification was submitted to ANOVA with the within-subjects factors of stimulus language (English versus French) and sentence type (trained versus novel) and the between-subjects factor of location (Montréal versus Storrs). The results showed a significant main effect of language, with talker identification higher for English compared to French [$F(1,37) = 174.05$, $p < .001$, $\eta_p^2 = 0.83$]. There was also a main effect of location: Montréal residents had higher accuracy in the test phase compared to the Storrs residents [$F(1,37) = 4.43$, $p = .042$, $\eta_p^2 = 0.11$]. Critically, there was an interaction between location and stimulus language [$F(1,37) = 12.47$, $p = .001$, $\eta_p^2 = 0.25$]. Planned comparisons showed that the two groups did not differ in their performance in English [$t(37) = -.10$, $p > .250$, $d = -0.03$], but instead differed in their performance in French [$t(37) = 3.07$, $p = .004$, $d = 0.98$], with Montréal residents outperforming the Storrs residents. This is striking when considering performance during the training phase, where – as a consequence of reaching learning criterion in fewer training blocks – Montréal residents received far less exposure to the French stimuli compared to Storrs residents. This pattern shows that incidental exposure to a nonnative language not only promotes talker learning for that nonnative language, it also facilitates memory for talkers' voices.

The ANOVA also revealed a main effect of trial type, with overall performance better for trained compared to novel sentences [$F(1,37) = 29.05$, $p < .001$, $\eta_p^2 = 0.44$]. However, there was a significant interaction between language and sentence type [$F(1,38) = 43.07$, $p < .001$, $\eta_p^2 = 0.54$]. Paired t -tests confirmed that there was no difference in performance for the trained and novel sentences in English [$t(38) = -.252$, $p > .250$, $d = -0.08$], but that there was a difference between the two sentence types in French [$t(38) = 6.80$, $p < .001$, $d = 2.21$], with talker identification for the trained items higher compared to performance for the novel items. There was no interaction between location and sentence type [$F(1,37) = 1.77$, $p = .192$, $\eta_p^2 = 0.05$], nor a three-way interaction [$F(1,37) = 2.24$, $p = .143$, $\eta_p^2 = 0.06$].

Recall that all 20 participants in the Montréal group and roughly half (8 out of 19) in the Storrs group had received formal classroom instruction in French. We did not expect this experience to influence talker identification performance in French, given that

none of the participants could comprehend the French sentences used in our study. Nonetheless, to determine whether differences in classroom instruction underlie our group differences, we repeated the analyses of training and test performance outlined above after removing the 11 subjects in the Storrs groups who did not have classroom French instruction. Despite the reduced power associated with a smaller sample, the interactions between location and stimulus language were numerically maintained for both the training and test data. The interactions described above did not reach statistical significance for the training data [$F(1,26) = 2.64$, $p = .116$, $\eta_p^2 = 0.09$], but did for the test data [$F(1,26) = 4.57$, $p = .042$, $\eta_p^2 = 0.15$]. These results show that differences in formal French instruction do not account for the observed differences in talker learning between our groups.

4. Discussion

Recent studies suggest that the native language benefit for talker identification reflects phonological knowledge and processing (Bregman & Creel, 2014; Perrachione et al., 2011). The results of our experiment support this as well: for both groups of English monolinguals tested here, learning new voices was easier in English than in French. However, our findings also indicate that neither functional phonological processing nor speaking proficiency is needed to promote a native language benefit for talker identification: English monolinguals with regular exposure to spoken French showed improved talker learning for French voices compared to English monolinguals without exposure to French. These results show that systematic exposure to a foreign language is sufficient for improving talker-learning abilities, even in the absence of functional language comprehension or speaking proficiency.

Our findings raise several important issues. First, it is unclear what level of language sensitivity or knowledge is needed to boost talker learning. The Montréal participants did not receive systematic exposure to French prior to adulthood, indicating that this is a late-learning effect. One possibility is that, through French exposure, Montréal participants have gained sensitivity to the structure of spoken French – including sensitivity to the phonetic, phonotactic, and prosodic features of French. Indeed, developmental research shows that infants can extract these properties from exposure to their native language even before word comprehension (e.g., Werker & Polka, 1993). Exposure to a wide range of talker variability may have accelerated the development of this structural sensitivity in the Montréal group. An alternative explanation for these effects is that Montréal participants may have a small lexicon of French words (e.g. greetings, food, street names),

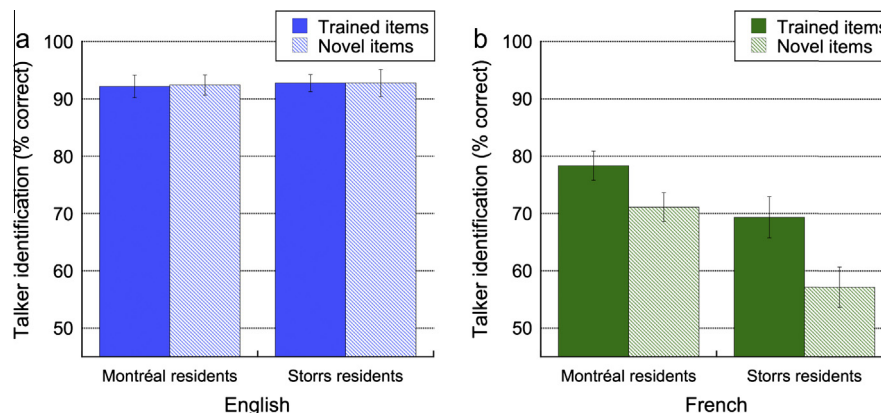


Fig. 4. Mean percent correct talker identification for trained and novel items for residents of Montréal and Storrs. Panel (a) shows test performance for English and panel (b) shows test performance for French. Error bars indicate standard error of the mean.

which may have been learned through basic French education or media exposure. Given that rudimentary phonological knowledge can be derived from a small lexicon (Pierrehumbert, 2001), it may be that the small lexicon gains are driving the benefit for talker identification. However, given that none of the participants could converse in French or understand the French sentences presented in the experiment, online comprehension during the experiment cannot explain differences in performance between the two groups.

Second, it is not known which aspects of language exposure drive these effects. While the major difference between the two groups is that participants from Montréal receive regular exposure to spoken French (and participants from Connecticut do not), it is also likely that Montréal participants are exposed to more French-accented English. Attentively listening to French-accented English may allow listeners to extract characteristics of the French language that are beneficial for learning French voices. Future work is needed to uncover which (or both) of these are responsible for these effects.

In any case, how might acoustic–phonetic knowledge improve talker-learning skills? Languages differ with respect to the cues that carry information about talker identity (Johnson, 2005). These cross-linguistic differences may bias speakers of different languages to attend to different cues in identifying talkers and lead to difficulty in tracking voices in a novel language (Bregman & Creel, 2014). Accordingly, exposure to an unfamiliar language spoken by multiple talkers may allow the perceiver to discover these optimal talker-distinguishing cues even without comprehending the language. In other words, Montréal participants may have gained sensitivity to the individual idiosyncrasies present in the productions of French phonological units; and in turn, may allow them to exploit these sensitivities for talker learning (Remez, 2010). Such perceptual learning may underlie the mechanisms that promote links between language exposure, language comprehension, and talker identification.

Recent infant research points to connections between linguistic experience and talker learning early in development, which suggest that talker learning precedes and bootstraps more sophisticated language processing skills. Specifically, 7-month-old infants show a native language advantage for discriminating talkers, when they are also becoming attuned to the surface phonetic properties of their native language (Johnson et al., 2011). Critically, both skills are gained through systematic language exposure that precedes word comprehension. Our findings show that, like infants, adults exposed to an unfamiliar language are also able to improve their talker identification skills prior to acquiring the phonological knowledge that is necessary to decode the acoustic signal and recognize words efficiently. Our findings are also consistent with prior research showing gradient effects of language experience on talker learning (Bregman & Creel, 2014): exposure to French provided a boost to talker identification in French, but does not seem to equate to native-like performance. Thus, our findings expand the language-familiarity gradient to include perceptual learning that precedes word comprehension.

Although the nature and length of language exposure required to boost talker learning is unclear, the present findings suggest that sustained, multi-talker exposure may be necessary. Important next steps for research include identifying the parameters of exposure that drive talker learning, clarifying the precise relationship

between phonetic sensitivity and talker learning, and exploring how advances in talker identification and phonetic perception support the development of abstract phonological processing. The present findings demonstrate that, across the life span, language exposure promotes processing benefits for that language, with talker identification benefits emerging before language comprehension. This finding raises the possibility that talker identification may bootstrap language acquisition, even for nonnative languages acquired in adulthood.

Acknowledgments

This research was supported by a Natural Sciences and Engineering Research Council Doctoral Scholarship to A. Orena, by a Social Sciences and Humanities Research Council grant (410-2006-1841) to L. Polka, and by faculty start-up funds (University of Connecticut, College of Liberal Arts and Sciences) to R. Theodore. Gratitude is extended to Terry Gottfried for helpful discussion, to Nicholas Monto and MaryKate Bisailon for their assistance with experiment programming, and to Anush Nersisyan, Alexina Hicks, and John Gerrity, Jr. for their assistance with subject testing.

References

- Bregman, M. R., & Creel, S. C. (2014). Gradient language dominance affects talker learning. *Cognition*, 130(1), 85–95.
- Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, 116, 3647–3658.
- Creel, S. C., & Bregman, M. R. (2011). How talker identity relates to language processing. *Language and Linguistics Compass*, 5, 190–204.
- DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science*, 208(4448), 1174–1176.
- Goggin, J. P., Thompson, C. P., Strube, G., & Simental, L. R. (1991). The role of language familiarity in voice identification. *Memory & Cognition*, 19(5), 448–458.
- Johnson, K. (2005). Speaker normalization in speech perception. In R. E. Remez & D. B. Pisoni (Eds.), *The handbook of speech perception* (pp. 363–389). Oxford: Blackwell Publishing.
- Johnson, E. K., Westrek, E., Nazzi, T., & Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science*, 14(5), 1002–1011.
- Köster, O., & Schiller, N. O. (1997). Different influences of the native language of a listener on speaker recognition. *Forensic Linguistics*, 4, 18–28.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940–967.
- Nygaard, L. C., Sommers, M. C., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5, 42–46.
- Perrachione, T. K., Del Tufo, S. N., & Gabrieli, J. D. E. (2011). Human voice recognition depends on language ability. *Science*, 333(6042), 595.
- Pierrehumbert, J. (2001). Why phonological constraints are so coarse-grained. *Language and Cognitive Processes*, 16(5), 691–698.
- Remez, R. E. (2010). Spoken expression of individual identity and the listener. In E. Morsella (Ed.), *Expressing oneself/expressing one's self: Communication, cognition, language, and identity* (pp. 167–181). New York: Psychology Press.
- Sullivan, K. P. H., & Schlichting, F. (2000). Speaker discrimination in a foreign language: First language environment, second language learners. *Speech, Language, and the Law*, 7, 95–111.
- Theodore, R. M., & Miller, J. L. (2010). Characteristics of listener sensitivity to talker-specific phonetic detail. *Journal of the Acoustical Society of America*, 128, 2090–2099.
- Valji, A. (2004). *Language preference in monolingual and bilingual infants*. Unpublished Master's thesis, McGill University, Montréal, Québec.
- Werker, J. F., & Polka, L. (1993). Developmental changes in speech perception: New challenges and new directions. *Journal of Phonetics*, 21, 83–101.
- Winters, S. J., Levi, S. V., & Pisoni, D. B. (2008). Identification and discrimination of bilingual talkers across languages. *Journal of the Acoustical Society of America*, 123, 4524–4538.