GRADIENT EFFECTS OF READING ABILITY ON NATIVE AND NON-NATIVE TALKER IDENTIFICATION

Minal A. Kadam[1]*, Adriel John Orena*[2]*, Rachel M. Theodore*[1]*, and Linda Polka*[2]

[1]University of Connecticut, [2]McGill University
rachel.theodore@uconn.edu

ABSTRACT

Talker identification is heightened for native compared to non-native talkers, suggesting that language comprehension influences talker recognition. Research has revealed a gradient effect of language experience on talker recognition; adults with reading disability show poor talker identification even in their native language. We examined whether this gradient would be observed among adults without reading disability. Monolingual English adults were assigned to the high or low reading group based on standardized assessments. All learned to identify the voices of English and French talkers and were then tested for retention of learning. The results indicate that compared to the low reading group, the high reading group (1) showed increased talker identification during training for both the native and non-native voices and (2) showed increased retention of learning, but only for the non-native voices. These results extend gradient effects of language proficiency on talker identification to include within-normal differences in reading ability.

Keywords: speech perception, talker identification, perceptual learning, reading disability

1. INTRODUCTION

The literatures on speech perception and talker recognition are historically distinct, reflecting the long-standing view that separate aspects of the speech signal are used to cue meaning and talker identity. This view is challenged by recent findings indicating that these aspects of the signal are fundamentally interviewed in the course of spoken language processing. With respect to speech perception, numerous findings indicate that experience with a talker's voice facilitates speech perception [9] and word recognition [4]. With respect to talker recognition, listeners are better able to identify talkers of their native language compared to a non-native language [2, 3, 7], which has been taken as evidence that phonological ability, knowledge of the sound structure of language, is an important mediator of talker recognition. Given these findings, a complete model of spoken language processing must describe how listeners integrate these two sources of information in the course of language comprehension.

Findings to date indicate that language ability exerts a gradient influence on native-language talker identification. Bregman & Creel [1] tested monolingual, native English listeners and bilingual Korean-English listeners on talker identification for English and Korean voices. Their results showed that talker identification was better in the respective native language for both groups of listeners. In addition, the Korean bilingual listeners performed better with the English voices than the monolingual English listeners performed with the Korean voices. Moreover, within the Korean-English bilingual listeners, those with more experience with the language (early English learners) had better talker identification performance compared to those with less experience (late English learners). These results provide strong evidence that the detriment listeners experience when identifying talkers of a non-native language is not absolute; rather, it reflects a continuum of experience and expertise in that non-native language.

Perrachione et al. [6] provided further evidence in support of a gradient effect of language competence on talker identification. In their study, adults with and without developmental dyslexia performed a talker identification task. Developmental dyslexia is a neurobiological disorder that leads to difficulties with fluent word decoding and reading ability [8]. The results showed that both groups of readers had poor talker identification for the non-native talkers, but the adults with reading disability also had poor talker identification for the native talkers. For the adults with dyslexia, there was a significant, positive correlation between accuracy in native language talker identification and performance on standardized measures of phonological processing.

Collectively, these findings suggest that in moving towards a model of spoken language processing that accounts for links between talker recognition and speech perception abilities, stability at the phonological level of processing may prove to be an important factor. However, future research is needed in order to determine whether phonological processing as captured in reading ability influences

talker identification not just for individuals with reading disability, but also across the range of values that comprise unimpaired variation.

To this end, the goal of the current work is to examine whether reading ability influences talker identification among typical readers. If the poor talker identification performance for native talkers observed in previous studies is the consequence of a dyslexia-specific neurobiological deficit, then we predict that it will not be observed for poor readers who perform above pathological thresholds. If, however, the poor talker identification performance reflects a gradient influence of phonological ability, then we predict that it will be observed across the typical spectrum of reading ability.

2. METHOD

2.1. Participants

Twenty-six monolingual speakers of American English with no history of speech, language, hearing, or neurological disorders were recruited for participation. Each completed a standardized diagnostic battery to assess reading sub-skills, reading comprehension, nonverbal intelligence, and working memory. The diagnostic battery consisted of the following standardized assessments: *Rapid Automatized Naming/Rapid Alternating Stimulus Tests*; the Elision, Blending Words, and Nonword Repetition subtests of the *Comprehensive Test of Phonological Processing*; the Sight Word Efficiency and Phonemic Decoding Efficiency subtests of the *Test of Word Reading Efficiency*, the Word Identification, Word Attack, and Passage Comprehension subtests of the *Woodcock Reading Mastery Tests - Third Edition*; the *Test of Nonverbal Intelligence - Fourth Edition*; and the Auditory Memory Index on the *Wechsler Memory Scales - Fourth Edition*. All participants scored within normal limits for tests of nonverbal intelligence and working memory.

Participants were assigned to either the high reading group or the low reading group based on performance on the comprehension sub-skills components of the battery. Specifically, a composite score was calculated for each participant, defined as the mean percentile score across the individual components, and a median split was used to determine placement into the high and low reading groups. Across participants, the mean percentile for the high reading group was 80 (SD = 5) and the mean percentile for the low reading group was 63 (SD = 11). An independent-samples t-test confirmed that these two distributions were statistically distinct [$t(24) = 4.93$, $p < .001$]. Thus, the high and low reading groups fall near the upper and middle range of the normal distribution, respectively, with both groups performing above the threshold used for identifying reading disability.

2.2. Stimuli

Auditory stimuli consisted of 12 English sentences and 12 French sentences that were matched in number of syllables and are described in detail by Valji [10]. Four native female speakers of each language produced each of the sentences for the respective language. Acoustic analyses confirmed that talkers of the two languages were equally discriminable on the basis of sentence duration, fundamental frequency, and variation in fundamental frequency. Two of the sentences were used during familiarization. Five of the sentences of each language were used during training and test phases, as described below. The remaining five sentences were only presented during the test phase in order to examine generalization of learning.

Visual stimuli consisted of eight cartoon faces, one for each talker. The faces were designed to be equally discriminable across the two languages.

2.3. Procedure

All participants complete a familiarization phase, a training phase, and a test phase for each language. The phases were completed in this order and were blocked by language, with language order counterbalanced across participants. During familiarization, participants listened to two sentences from each of the four talkers as it was paired with one of the four cartoon faces in order to begin associating each voice with the appropriate cartoon face.

The training phase consisted of blocks of 60 randomized trials (4 talkers X 5 sentences X 3 repetitions). Each trial consisted of presentation of an auditory sentence produced by one talker and a visual array consisting of two cartoon avatars. Participants were directed to indicate which of the faces matched the voice. Feedback was provided on every training trial. Successive training blocks were completed until each participant met the learning criterion defined as 85% correct or higher in a single training block or the completion of eight training blocks.

The test phase consisted of 120 trials (4 talkers X 10 sentences X 3 repetitions). On each test trial, participants heard an auditory sentence produced by one of the four talkers and saw the four cartoon faces. They were directed to indicate which of the four faces matched the talker's voice. No feedback

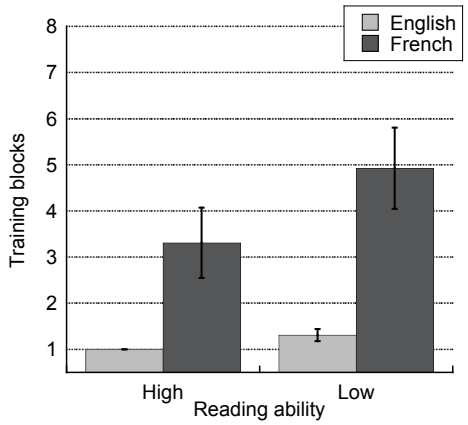
was provided at test. The entire procedure lasted approximately 90 minutes.

3. RESULTS

3.1. Training

Performance during training was quantified in three ways: (1) number of training blocks required to meet the learning criterion, (2) percent correct talker identification during the first training block, and (2) percent correct talker identification during the final training block. First we considered performance as indexed by the number of training blocks required to meet learning criterion, as shown in Figure 1. Error bars in all figures show standard error of the mean.

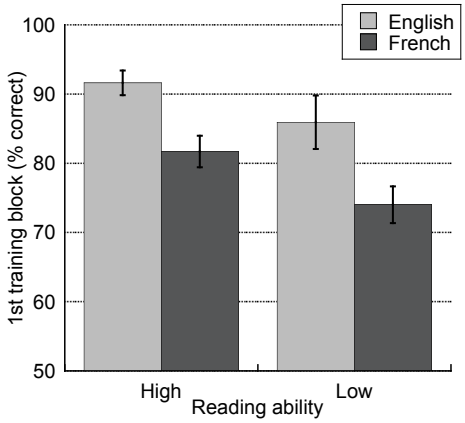
Figure 1: Number of training blocks required to meet learning criterion.



Visual inspection suggests that effect of reading ability on learning rate was negligible for the English talkers, but that there was a numerical difference for the French talkers, with the low reading group requiring more training blocks to reach criterion compared to the high reading group. These data were submitted to ANOVA with the between-subjects factor of reading ability (high versus low) and the within-subjects factor of language (English versus French). The results confirmed a main effect of language, with fewer blocks required for the English compared to the French talkers [$F(1,24) = 25.83$, $p < .001$]. However, the main effect of reading ability did not reach statistical significance [$F(1,24) = 2.66$, $p = .116$], nor was there an interaction between language and reading ability [$F(1,24) = 1.26$, $p = .273$].

Next we considered performance during the first training block for each language, which is shown in Figure 2. Mean percent correct talker identification was submitted to ANOVA with reading ability and language as factors. As expected, the results of the ANOVA showed a robust main effect of language [$F(1,24) = 15.63$, $p = .001$], with accuracy in the first block higher for the native compared to non-native talkers. In addition, there was a main effect of reading ability [$F(1,24) = 6.01$, $p = .022$], with those in the high reading group showing increased accuracy compared to the low reading group for both the English and French voices. There was no interaction between language and reading ability [$F(1,24) = 0.14$, $p = .716$]. These results indicate that reading ability influenced talker identification during the first block of training, with better performance for those in the high compared to the low reading group for both the native and non-native talkers.

Figure 2: Percent correct talker identification for the first training block.

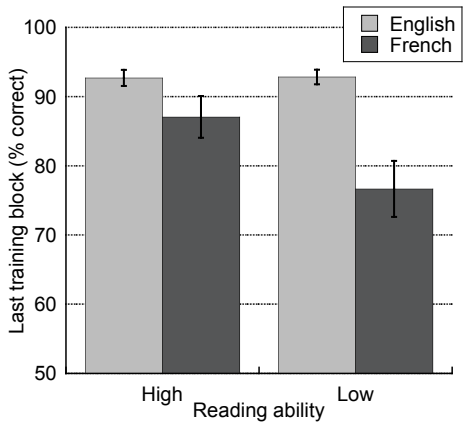


The final training analysis considered performance during the last block of training, which is shown in Figure 3. These values were submitted to ANOVA following the structure outlined previously. The ANOVA revealed a main of language [$F(1,24) = 20.24$, $p < .001$], with percent correct talker identification higher for the native compared to the non-native voices. There was a marginal main effect of reading ability [$F(1,24) = 3.23$, $p = .085$], and a statistically significant interaction between language and reading ability [$F(1,24) = 4.71$, $p = .040$]. Independent t-tests showed that accuracy was increased for the high compared to the low reading group for the French voices [$t(24) = 2.06$, $p = .051$], but not for the English voices [$t(24) = 0.82$, $p = .935$].

Collectively, the results from the training data indicate that when considering performance in terms of percent correct talker identification, reading ability systematically influenced performance. At the beginning of the training period, performance for both the native and non-native talkers was decreased in the low compared to the high reading group. At the end of training, this pattern continued but only for the non-native talkers. Neither of these patterns were captured when learning was measured at the

more gross level of number of training blocks.

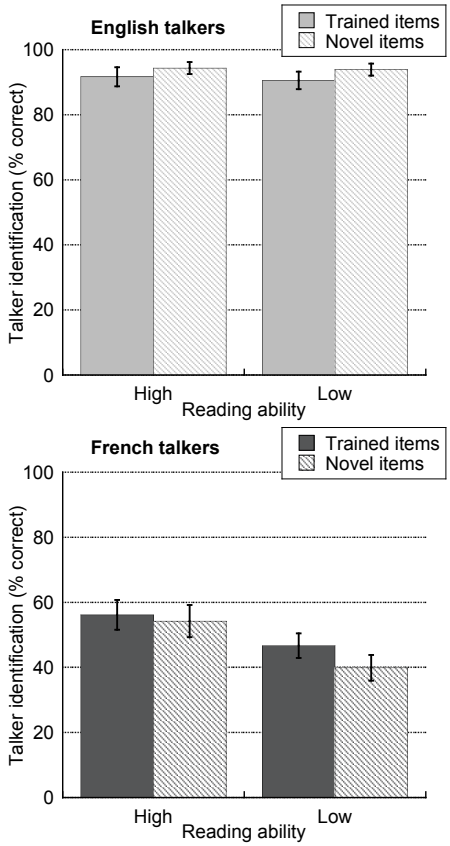
Figure 3: Percent correct talker identification for the last training block.



3.2. Test

Performance at test was measured in terms of percent correct talker identification, which was calculated for each participant separately for the English and French voices and for the trained and novel sentences. Figure 4 shows mean accuracy for the high and low reading groups for the native (top panel) and non-native voices (bottom panel).

Figure 4: Percent correct talker identification at test for the English and French voices.



Percent correct talker identification was submitted to ANOVA with the factors of reading ability, language, and sentence type (trained versus novel). The results showed a main effect of language, with higher talker identification accuracy for the native compared to the non-native talkers [$F(1,24) = 370.86$, $p < .001$]. There was also an interaction between language and sentence type [$F(1,24) = 11.44$, $p = .002$], with talker identification more accurate for the trained compared to the novel sentences for the French voices [$t(25) = 2.10$, $p < .05$], but slightly higher for the novel compared to the trained sentences for the English voices [$t(25) = 3.01$, $p < .01$], reflecting increased generalization for the native compared to the non-native voices. There was also a significant interaction between language and reading ability [$F(1,24) = 6.06$, $p = .021$]. T-tests showed that the two reading groups did not differ for the English voices [$t(24) = 0.26$, $p = .798$], but did for the French voices [$t(24) = 2.06$, $p = .051$], with performance improved for the high compared to the low reading group. There was no main effect of trial type or reading ability, no interaction between trial type and group, and the 3-way interaction was not reliable ($p > .130$ in all cases).

4. CONCLUSION

There is a growing body of evidence indicating not only that listeners integrate talker identity and linguistic content, but that these aspects can mutually inform and constrain each other. Research suggests that language ability, as measured in terms of stability in phonological processing, exerts a gradient influence on talker identification [2, 6]. The results here are consistent with this account. Adults with reading ability near the top of the normal distribution showed heightened talker identification compared to adults with reading ability near the middle of the distribution. This finding extends earlier work showing impaired talker recognition in adults with dyslexia to include a gradient influence of phonological processing on talker identification even within the unimpaired range of reading ability.

Though phonological stability influences talker identification, recent findings suggest that this is not the sole determinant. Native-language benefits are observed in the absence of language comprehension [2, 3], and mere exposure to a non-native language can improve talker recognition in that language [5]. In moving towards a model of spoken language processing that describes the interplay between language comprehension and talker recognition, the contributions of language exposure and expertise must be delineated across the developmental trajectory. Future work is aimed at this goal.

5. REFERENCES

[1] Bregman, M. R., Creel, S. C. (2014). Gradient language dominance affects talker learning. *Cognition*, *130*, 85-95.

[2] Fleming, D. Giordano, B. L., Caldara, R., Belin, P. (2014). A language-familiarity effect for speaker discrimination without comprehension. *Proceedings of the National Academic of Sciences, 111,* 13795-13798.

[3] Johnson, E. K., Westrek, E., Nazzi, T., Cutler, A. (2011). Infant ability to tell voices apart rests on language experience. *Developmental Science, 14,* 1002-1011.

[4] Nygaard, L. C., Sommers, M. S., Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science, 5,* 42-46.

[5] Orena, A. J., Theodore, R. M., & Polka, L. Exposure is all it takes to improve talker learning in a foreign language. Submitted.

[6] Perrachione, T. K., Del Tufo, S. N., Gabrieli, J. D. (2011). Human voice recognition depends on language ability. *Science, 333,* 595-595.

[7] Perrachione, T. K. Wong, P. C. (2007). Learning to recognize speakers of a non-native language: implications for the functional organization of human auditory cortex. *Neuropsychologia, 45,* 1899-1910.

[8] Pugh, K. R., Mencl, W. E., Jenner, A. R., Katz, L., Frost, S. J., Lee, J. R., Shaywitz, S. E., Shaywitz, B. A. (2000). Functional neuroimaging studies of reading and reading disability (developmental dyslexia). *Mental Retardation and Developmental Disabilities Research Reviews*, *6*, 207-213.

[9] 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.

[10] Valji, A. (2004). Language preference in monolingual and bilingual infants. Unpublished Master's thesis, McGill University, Montréal, Québec.