Research Article

Language Ability and the Familiar Talker Advantage: Generalizing to Unfamiliar Talkers Is What Matters

Susannah V. Levi,^a Daphna Harel,^b and Richard G. Schwartz^c

Purpose: Previous studies with children and adults have demonstrated a *familiar talker advantage*—better word recognition for familiar talkers. The goal of the current study was to test whether this phenomenon is modulated by a child's language ability.

Method: Sixty children with a range of language ability were trained to learn the voices of 3 foreign-accented, German–English bilingual talkers and received feedback about their performance. Both before and after this talker voice training, children completed a spoken word recognition task in which they heard consonant–vowel–consonant words mixed with noise that were spoken by the 3 familiarized talkers and by 3 unfamiliar German–English bilinguals.

Results: Two findings emerged from this study: First, children with both higher and lower language ability performed similarly on the familiarized talkers. Second,

children with higher language scores performed similarly on both the familiarized and unfamiliar talkers, whereas children with lower language scores performed worse on the unfamiliar talkers compared to familiar talkers, suggesting an inability to generalize to novel, unfamiliar talkers who spoke with a similar accent.

Discussion: Together, these findings indicate that children with higher language scores are able to generalize knowledge about foreign-accented talkers to help spoken word recognition for novel talkers with the same accent. In contrast, children with lower language skills did not exhibit the same magnitude of generalization. This lack of generalization to similar talkers may mean that children with lower language skills are at a disadvantage in spoken language tasks because they are unable to process speech as well when listening to unfamiliar talkers.

The speech signal is highly variable, due in part to differences across talkers, who vary in many ways, including gender, size, dialect or nativeness status, and idiolect. Many studies have shown that listeners process speech in a talker-contingent manner, where perception depends on who is talking (Allen & Miller, 2004; Eisner & McQueen, 2005; Kraljic & Samuel, 2007; Ladefoged & Broadbent, 1957; McQueen, Cutler, & Norris, 2006; Norris, McQueen, & Cutler, 2003; Theodore & Miller, 2010). Research on the *familiar talker advantage* (FTA)—where listeners are better at understanding speech spoken by familiar

talkers than unfamiliar talkers—shows that talker-contingent processing can result not only in a change in perception but also in an improvement in perception in listeners who have no known language impairments. The primary goal of the current study is to expand on previous research on the FTA in children by exploring whether a child's language ability modulates performance. As the talkers in this study are all German–English bilinguals, a secondary goal is to test whether language ability modulates a child's ability to generalize learned information about a particular type of foreign-accented speech to a novel set of similarly accented talkers. The results of this study will determine whether children with poorer language skills also successfully integrate information about a talker to improve spoken language processing.

^aDepartment of Communicative Sciences and Disorders, New York University, New York

^bPRIISM Applied Statistics Center, Department of Applied Statistics, Social Science, and Humanities, New York University, New York

^cProgram in Speech-Language-Hearing Sciences, Graduate Center, City University of New York, New York

Correspondence to Susannah V. Levi: svlevi@nyu.edu

Editor-in-Chief: Sean Redmond

Editor: Lisa Archibald

Received April 26, 2018

Revision received October 8, 2018

Accepted January 8, 2019

https://doi.org/10.1044/2019_JSLHR-L-18-0160

FTA, Speech Perception, and Language Disorders

The primary goal of the current study is to establish whether children with both typical and poorer language ability show the FTA. Previous studies have found evidence for an FTA across the life span, from school-age

Disclosure: The authors have declared that no competing interests existed at the time of publication.

children (Levi, 2015) to young adults (Levi, Winters, & Pisoni, 2011; Nygaard & Pisoni, 1998; Nygaard, Sommers, & Pisoni, 1994; Yonan & Sommers, 2000) to older adults (Yonan & Sommers, 2000), suggesting that it is a robust effect. Previous work on the FTA, however, has also documented limitations to the FTA. For example, the FTA is found only for listeners who learned the talkers sufficiently well during training (greater than 70% accuracy during training; Nygaard & Pisoni, 1998; Nygaard et al., 1994). In addition, a recent study using an implicit voice-learning task, as opposed to an explicit voice-learning task, suggests that short-term implicit learning is not sufficient to generate the FTA (Case, Seyfarth, & Levi, 2018). Thus, although the FTA has been found across many studies, it is important to test limitations on this phenomenon to understand in which situations and for what types of listeners talker familiarity can aid spoken language processing.

One important consideration to examine is the contribution of a participant's language ability. All of the previous studies of the FTA have tested individuals with typical language ability. Research on children with language disorders—referred to as developmental language disorder or specific language impairment—suggests that speech perception and auditory processing in these children are also impaired (McArthur & Bishop, 2004; McReynolds, 1966; Rosen, 2003; Shafer, Morr, Datta, Kurtzberg, & Schwartz, 2005; Stark & Heinz, 1996; Sussman, 1993, 2001; Tallal & Piercy, 1974, 1975; Tallal & Stark, 1981; Wright et al., 1997). In these previous studies, the tasks used to assess auditory perceptual abilities tend to be relatively simple (e.g., discrimination) and to test a small number of phonemes (e.g., vowel contrasts) or auditory skills (e.g., tone detection). For example, Wright et al. (1997) and McArthur and Bishop (2004) examined tone detection with different types of maskers. In terms of phonetic perception, several studies have examined discrimination and identification of vowel contrasts (Shafer et al., 2005; Sussman, 1993, 2001). In contrast to these previous studies, the current study examines spoken language processing in a more complex yet more natural task (spoken word recognition) and thus tests changes in performance across the range of phonemes in English. The task of spoken word recognition represents an experimental task that is a more relevant and natural type of speech perception task in that it examines perception of whole words.

To have an FTA, listeners must (a) store information about a talker and then (b) use that information when completing a linguistic perception task (such as spoken word recognition). To store information about different talkers, listeners must be able to, at a minimum, discriminate between the voices of different talkers. Two studies have examined whether language ability affects how well children process information about different talkers' voices. Levi and Schwartz (2013) tested school-age children on a talker discrimination task where listeners heard two words and were asked to determine whether the words were spoken by the same talker or two different talkers. No differences

were found between children with typical language development and those with language impairment, although the sample size was small. In contrast, Dailey, Plante, and Vance (2013) tested preschool children on a talker discrimination task and found poorer performance for children with language impairment compared to their typically developing peers. If children with language impairments have poorer talker discrimination ability, it is possible that these children will be unable to adequately store information about different talkers and then use that to improve speech perception and spoken language processing.

Alternatively, although the results of Dailey et al. (2013) suggest an impairment in talker discrimination for children with language impairments, it is possible that talker discrimination ability, though worse than for children with typical language development, is still sufficient to result in an FTA. This would be an important contribution, as it would provide a possible mechanism to improve processing and possibly free up cognitive resources that could be used for higher level processing. Although not a direct test of whether language ability modulates the FTA, two previous studies found that individuals with poor baseline spoken language processing benefit from talker familiarity more than those with better baseline spoken language processing. Markham and Hazan (2004) examined online attunement to a talker's voice by comparing performance in two conditions: (a) where listeners heard a word and repeated it and (b) where they heard the carrier phrase "and now please say" followed by three words in the same voice. For listeners with baseline performance in the bottom quartile, providing the carrier phrase resulted in a significant increase in accuracy. The authors argued that brief exposure to a talker's voice might be the most beneficial for listeners with the poorest spoken language processing skills. Similarly, in a study of the FTA, Levi (2015) found that typically developing children with the lowest level of baseline spoken word recognition accuracy showed the most benefit (biggest difference) for talker familiarity in a talker training FTA study. Thus, it is possible that children with language disorders will also benefit from familiarity to a talker's voice.

In summary, previous studies have found that listeners across a range of ages can understand speech better when spoken by a familiar talker. The current study extends these findings by examining whether children with poorer language skills, who have some documented perceptual impairments in both the linguistic domain (e.g., phoneme perception) and the talker domain (e.g., talker discrimination), are also able to harness talker information as a compensatory strategy to improve spoken word recognition. Whereas previous studies of the FTA either do not explicitly control for or report language ability (e.g., Levi et al., 2011; Nygaard & Pisoni, 1998) or only test participants whose language skills are within the normal range (Levi, 2015), the current study explicitly tests the role of a participant's language ability in modulating the FTA.

Adaptation to Foreign Accents

A second goal of the current study is to test generalization of learned talker characteristics of one set of foreign-accented talkers to novel talkers with the same foreign accent. Several studies have shown that listeners are able to adapt to foreign-accented speech when presented with speech from multiple talkers with the same foreign accent. Bradlow and Bent (2008) tested adult listeners' ability to generalize knowledge of Chinese-accented talkers to a novel Chinese-accented talker. The training paradigm consisted of a spoken sentence recognition task with different sets of talkers. Listeners who completed the training portion with multiple (five) Chinese-accented talkers outperformed those who had been exposed to only a single Chinese-accented talker, suggesting that exposure to a range of talkers allowed the listeners to generalize this particular foreign accent. Similar generalization results were found by Sidaras, Alexander, and Nygaard (2009). In their study, listeners exposed to multiple (six) Spanish-accented talkers showed better word and sentence recognition accuracy for novel Spanish-accented speech. As with Bradlow and Bent (2008), these results show that listeners are able to learn accent-general information when exposed to more than one talker with that accent type. Indeed, Bradlow and Bent (2008) found no learning benefits when listeners were presented with a Slovakian-accented talker.

In contrast to these studies of accent-general learning, a previous study in our laboratory with a smaller set of participants than the current study did not find any accent-general learning. Levi (2015) trained children to learn the voices of three German–English bilingual talkers. Before and after training, children completed a single-word recognition task where they heard words produced by both these trained/familiarized talkers and by three additional German–English bilingual talkers. In that study, no accent-general learning was found; performance at baseline and post-training did not differ for the untrained talkers. In that study, lack of accent-general adaptation was attributed to possible differences in the training task; Bradlow and Bent (2008) and Sidaras et al. (2009) examined generalization where training and test both targeted linguistic processing (transcription), whereas in Levi (2015), training was in the form of voice identification, and this did not transfer to improvement for linguistic processing (spoken word recognition) for novel talkers with the same foreign accent.

The lack of accent-general learning in the smaller Levi (2015) study was surprising given the studies that have found accent-general learning in adults. The current study allows us to test the replicability of this null result by changing two aspects of the study. First, the current study tests a larger set of participants than the original study. Second, the current study tests the modulating effect of language ability, which was not included as a predictor in the analysis of the original, smaller data set. Both of these factors—smaller sample size and treating all children whose language ability was greater than 85 on a standardized test of

language ability as a single group—could have contributed to the null results in the original study.

Current Study

The current study builds on previous research in our laboratory on the FTA and examines performance by a group of children with a wider range of language abilities than the typically developing children included in Levi (2015). The current study directly explores whether language ability modulates the FTA in school-age children. It builds on previous research showing that those individuals with the poorest spoken language processing skills show the most benefit from talker familiarity. Although possible that individuals with the poorest speech perception are those with poorer language ability, the current study offers a direct test of the contribution of language skills on the integration of talker information for linguistic processing. As all talkers in the study are German–English bilinguals, the baseline posttraining design allows us to test whether familiarity with German-accented English generalizes to improved perception of German-accented speech produced by novel talkers.

Method

Participants

Sixty children ages 6;8–12;0 (years;months; $M = 9;3$) participated in the study, with 35 from the original (Levi, 2015) data set. All children were native speakers of American English, with English as their dominant language as indicated through parent report. No child had exposure to German or German-accented English. All children completed the subtests of the Clinical Evaluation of Language Fundamentals–Fourth Edition (CELF-4; Semel, Wiig, & Secord, 2003) that generate the Core Language composite score to assess language ability. This score is normed at 100, with an SD of 15. The Core Language composite scores for the 60 children ranged from 54 to 133 ($M = 98.4$). In addition, children completed the Test of Nonverbal Intelligence–Third Edition (Brown, Sherbenou, & Johnsen, 1997) to assess nonverbal cognition. Like the CELF-4, this test is normed at 100, with an SD of 15. Test of Nonverbal Intelligence–Third Edition scores ranged from 70 to 146 ($M = 105.6$). Pearson correlations indicated that these two scores were highly correlated ($r = .735$, $p < .001$). Thirty-seven children passed a hearing screening at 25 dB HL at 1000, 2000, and 4000 Hz. Hearing for the remaining 23 children was determined as follows: For nine children, the parent reported that the child passed a hearing screening within 12 months prior to participation, and for 14 children, the parent did not indicate any known hearing problems.

Stimuli

Stimuli were drawn from a database of German–English bilingual talkers producing consonant–vowel–consonant

words in both English and German. All stimuli were recorded in a sound-attenuated booth using a SHURE 98 head-mounted microphone and digitized onto a Tucker-Davis Technologies System II. See Levi, Winters, and Pisoni (2007) and Levi et al. (2011) for more information about recording. For the current study, only English stimuli from six female talkers were used. These talkers were selected to have a similar level of intelligibility but had varied fundamental frequencies to increase the likelihood that children would identify the talkers during training (see Table 1).

The words used in the current study were rated independently by two speech-language pathologists and the first author as either likely to be known by children ages 5–8 years (“high lexical familiarity”) or likely to be unknown to these children (“low lexical familiarity”; see Appendix). Age of acquisition (Cortese & Khanna, 2008) for these two groups of words was also checked. Words in the high-familiarity group had a mean age of acquisition of 3.01 (acquired between ages 4 and 6 years), whereas words in the low-familiarity group had a mean of 5.34 (acquired after ages 8–10 years). Three of our high-familiarity words (*house*, *mouse*, *pill*) and 17 low-familiarity words were not found in Cortese and Khanna’s database. Based on Kučera and Francis (1967), 15 of 17 of these low-frequency words had a frequency of 1, and the remaining two words (*couth*, *goon*) were not in the database.

Twenty-four high- and 24 low-familiarity words were randomly selected for the baseline and posttraining word recognition tasks. From the remaining words, 56 high- and 56 low-familiarity words were randomly selected for the training task. Twelve additional high-familiarity words were used as practice in both word recognition tasks. This resulted in 116 distinct high-familiarity words (12 + 24 + 54 + 24) and 104 distinct low-familiarity words (24 + 54 + 24). Eight such randomly selected lists were created, and children were randomly assigned to each list. Both high- and low-familiarity words were included in the current study because it allows us to test whether perceptual improvement is limited to processing words with stronger lexical representations (high-familiarity words) compared

to those with weaker or no lexical representations (low-familiarity words). We return to this point in the discussion.

Stimuli in the two spoken word recognition tasks were mixed with signal-dependent noise based on Schroeder (1968) and Benkí (2003) and using MATLAB code from Felty (2007). Stimuli were mixed with noise to avoid ceiling effects during the baseline spoken word recognition task. The first 13 children included in the current study completed the spoken word recognition tasks with a signal-to-noise ratio (SNR) of either 10 dB ($n = 7$) or 15 dB ($n = 6$). Because several children performed well with this SNR, subsequent children completed the task with an SNR of +5 dB to avoid ceiling effects. Thus, SNR was included in the data analysis as a control variable in all analyses.

Procedure

Children completed three experimental tasks: baseline spoken word recognition, talker training, and posttraining spoken word recognition. Baseline spoken word recognition and Training Day 1 were completed on the same day. This was followed by four additional training days. On the sixth day, children completed the generalization task (no feedback) and the posttraining spoken word recognition task. Five days of talker voice training was selected because previous work with adults using these stimuli had found an effect with 3 days of training (Levi et al., 2011). Two training days were added to provide additional training in the event that children would need more training than adults. Experiments were administered with E-Prime 2.0 Professional (Schneider, Eschman, & Zuccolotto, 2007) and presented binaurally over Sennheiser HD-280 circumaural headphones. Children were tested in a quiet room in their school or home, or in the Department of Communicative Sciences and Disorders at New York University.

Baseline and Posttraining Spoken Word Recognition

Children completed two spoken word recognition tasks, one at baseline before talker training and one after training. In each task, children heard words mixed with

Table 1. Information on familiar and unfamiliar talkers.

Talker type	Talker	AoA	Years of English	Length of residence	English proficiency	F ₀ (Hz)	Intelligibility	Degree of foreign accent
Familiar talkers	F3	10	14	1	5	222	76.2%	2.73
	F8	13	16	4	5	170	83.3%	1.90
	F11	—	—	2	4.5	260	77.8%	4.04
Unfamiliar talkers	F4	13	13	3	4.5	186	79.4%	2.35
	F9	9	16	2	4	202	79.4%	2.51
	F12	7	26	5	5	186	78.6%	—

Note. Larger ratings indicate a higher degree of foreign accent (Talker F12 was not included in the study of degree of foreign accent). AoA = age of acquisition of English; Years of English = number of years speakers have been learning/using English (age at test – age of acquisition); Length of residence = years in the United States; English proficiency = self-reported measure (1 = *poor*, 5 = *fluent*); F₀ = average fundamental frequency at the vowel midpoint; Intelligibility = average percent correct of words correctly identified by adults in quiet (Levi et al., 2011); Degree of foreign accent = average foreign accent rating on a 7-point Likert scale (Levi et al., 2007).

noise and were asked to say the word they heard. They were told to repeat what they heard even if they did not know the word. First, they completed a practice block with 12 high-familiarity words. This was followed by two experimental blocks containing 24 high- and 24 low-familiarity words in random order. Half of the words were produced by three talkers who would be learned during training (*familiar talkers*), and half were produced by three additional talkers (*unfamiliar talkers*). There was no overlap in the words used for either of the spoken word recognition tasks or the training task. Responses were transcribed online by one phonetically trained researcher and also recorded and transcribed offline by a second researcher. A third transcriber resolved any discrepancies in the two transcriptions.

Talker Training

Talker training was completed over 5 days, typically with no more than 2 days between trainings. To check that there was no systematic relationship between language ability and the number of days taken to complete training, we conducted a correlation on the Core Language composite from the CELF-4 and the total days between baseline and post-training spoken word recognition. This correlation revealed neither strong nor significant ($r = -.153, p = .243$) associations between these two measures. On each day of training, children completed two identical training sessions (with feedback) and one test session (no feedback). In each training session, children first heard two high- and two low-familiarity words produced by each of the three talkers whose voices they would be learning. As the words were played, the image of the talker (represented by a cartoon-like animal) was presented on the screen. Following this brief exposure block, children were presented with 30 trials (5 high- and 5 low-familiarity words $\times$ 3 talkers) in which they heard a single item and were asked to click on the cartoon character who spoke that word. After children made their response, they received feedback regarding their accuracy in the form of a smiley or frowny face, and the word was played again with only the image of the correct talker on the screen. Items were presented in random order. The identical training session was completed twice. This was followed by a test session with different words that was identical to the training session, except that children did not receive feedback. After 5 days of training, on the sixth day, children completed a generalization task that was like the test session (no feedback) and included 24 old words that had been heard on Day 1 of training and 24 novel words. Thus, children were presented with 56 high- and 56 low-familiarity words (same two high-/low-familiarity items during exposure block, 5 high-/low-familiarity items $\times$ 5 days for the training sessions with feedback, 5 high-/low-familiarity items $\times$ 5 days for the test sessions without feedback, and four high-/low-familiarity items during generalization).

Statistical Analyses

Generalized logistic mixed-effects models were fit to the data using the lme4 (Bates, Mächler, Bolker, & Walker,

2015) package in RStudio Version 1.0.136 (R Development Core Team, 2017) and following Harel and McAllister (in press). The response variable in all analyses was whole-word accuracy. Separate analyses were conducted for the high-familiarity words coded for whole-word correct and low-familiarity words coded for whole-word correct. Models for high- and low-familiarity words were conducted separately to avoid the inclusion of four-way interaction terms in the model, thus increasing model interpretability. For each analysis, three models were considered.

The maximal model included fixed-effects terms for the percentage of words correctly recognized at baseline, SNR, age (in months), gender, talker type (familiar, unfamiliar), talker learning (percent correct talker identification on Days 4 and 5 of training), Core Language composite score, and all two- and three-way interactions between talker type, talker learning, and Core Language composite scores. The maximal model also included random intercepts for each talker and for each child to adequately model the repeated measures obtained from each of these two sources. Talker type and participant gender were coded through the sum coding procedure to aid in the interpretation of the main effects. The second model removed the three-way interaction between talker type, talker learning, and Core Language composite score but kept all two-way interactions. The third model further removed the two-way interaction term between Core Language composite score and talker learning. No further models were considered because both of the other two-way interaction terms were statistically significant in the third model for all analyses below. Model selection was performed using the Akaike information criteria (AIC; Akaike, 1974). The AIC is a penalized version of the log likelihood of the model, a measure of goodness of fit of the model to the data, which penalizes the log likelihood by a function of the number of parameters estimated by the model (i.e., the complexity of the model). Models with the lowest AIC are preferred.

Results

Spoken Word Recognition for High-Familiarity Words

For the analysis of high-familiarity whole-word correct, the AIC selected the third, simplest model as the best fitting model to the data.¹ This final model showed a negative slope for the percentage of words correctly

¹Note that the transcriptions of responses were also coded for phoneme accuracy. In this case, transcriptions of responses were matched to the target transcription to give participants maximal credit. For example, response *age* [eɪdʒ] for target *sage* [seɪdʒ] or response *close* [klaʊz] for target *loathe* [louð] were both coded as two phonemes correct. Participants could receive zero to three phonemes correct per trial. The same three generalized logistic mixed-effects models were fit to the high-familiarity phoneme data, where responses were coded for phonemes rather than whole-word correct. The AIC selected the same model for phoneme accuracy and word accuracy, which is unsurprising as these two measures were highly correlated ($r = .858$).

recognized at baseline ($\beta = -0.040$, $z = -2.829$, $p = .004$). SNR was statistically significant ($\beta = 0.09$, $z = 5.472$, $p < .001$) indicating that the higher the SNR, the better the child performed at the task. Age and participant gender were not statistically significant in the model, indicating that these variables had no association with performance. Core Language composite score was significant ($\beta = 0.249$, $z = 3.800$, $p = .001$), with a positive slope indicating that children with better language abilities recognized more words after training, on average. However, the interaction between talker type and Core Language composite score was also statistically significant. This indicates that the slope for Core Language composite score was actually more positive for unfamiliar talkers ($\beta = 0.095$, $z = 3.027$, $p = .002$) as compared with familiar talkers. As shown in the left panel of Figure 1, children with high Core Language scores performed similarly on both familiar and unfamiliar talkers, whereas children with low Core Language scores exhibited poorer performance only for the unfamiliar talkers, suggesting that these children did not generalize talker information from the training phase to improving spoken word recognition for novel talkers. The main effects of talker type and talker learning were not statistically significant, but the interaction of talker type and talker learning was statistically significant. This indicates that the slope for talker learning was only positive for unfamiliar talkers ($\beta = 0.08$, $z = 3.241$, $p = .001$). Children showed similar performance for familiar talkers regardless of how well they learned the talkers. In contrast, children who learned the (familiar) talkers well also performed well on the unfamiliar talkers, suggesting generalization to these novel talkers.

Spoken Word Recognition for Low-Familiarity Words

The three generalized logistic mixed-effects models were also fit to the low-familiarity word data. In contrast to the high-familiarity words, the AIC selected the maximal model that included all two- and three-way interactions between talker type, talker learning, and Core Language composite scores.²

This model showed a negative slope for the percentage of words correctly recognized at baseline ($\beta = -0.112$, $z = -8.454$, $p < .001$). SNR was statistically significant ($\beta = 0.076$, $z = 2.847$, $p = .004$), indicating that the higher the SNR, the better the child performed. Age and participant gender were not statistically significant in the model, indicating no performance differences across these two variables. The Core Language composite score was significant ($\beta = 0.226$, $z = 2.409$, $p = .016$), with a positive slope indicating that the better a child's overall language abilities, the more words he or she recognized after training. However, the interaction between talker type and Core Language composite score was also statistically significant.

²Again, the same generalized linear mixed-effects models were fit to the phoneme data, and similar results were found.

This indicates that the slope for the Core Language composite score was more positive for unfamiliar talkers ($\beta = 0.145$, $z = 5.589$, $p < .001$) than for familiar talkers, similar to the findings for the high-familiarity words. In other words, children with higher language scores performed similarly for the familiar and unfamiliar talkers, whereas children with lower language scores were disproportionately negatively affected when listening to unfamiliar talkers.

The main effects of talker type and talker learning were not statistically significant, but their interaction was statistically significant. This indicates that the slope for learning was only positive for unfamiliar talkers ($\beta = 0.121$, $z = 4.593$, $p < .001$). Furthermore, the three-way interaction between talker type, talker learning, and Core Language composite score was statistically significant ($\beta = 0.280$, $z = 12.275$, $p < .001$), indicating that children who both learned the talkers well during training (high scores on talker training) and had high Core Language composite scores saw an additional increase in their performance on unfamiliar talkers.

Discussion

The current study addressed two main questions: (a) whether language ability modulates perception of familiarized talkers and (b) whether it modulates perception of novel talkers who have a similar foreign accent. Results revealed a larger difference in performance between children with poorer versus better language skills for the unfamiliar talkers (steeper slope) than the familiar talkers. As will be discussed further below, this suggests that performance was more similar between children with poorer versus better language skills on the familiar talkers. In addition to our primary findings, we also found a relationship between how well children learned the talkers' voices and how they performed on the spoken word recognition task.

Language Ability, the FTA, and Accent-General Learning

Previous research on the FTA did not consider whether a participant's language skills would affect their ability to learn the voices of a set of talkers and then use that stored talker information to improve spoken word recognition for those talkers. The current results suggest that language ability may not modulate the benefit that listeners show for familiar talkers but does modulate whether listeners generalize knowledge of a learned type of talker (here, German-accented talkers) to novel talkers with the same accent. This is both a promising and disconcerting finding. These data suggest that children with poorer language skills are able to harness stored knowledge of talker-specific production information to enhance speech perception. This is a positive finding and suggests that children with poorer language skills may also use this information to aid in auditory stream segregation for

Figure 1. Mean word recognition accuracy (by participant) after training as a function of Core Language composite score (left) and talker learning (right), shown separately for familiar (black dashed line) and unfamiliar (solid gray line) talkers.

familiar talkers (e.g., teachers) in noisy environments (e.g., classrooms), as talker familiarity has been shown to aid auditory stream segregation across a range of ages (Barker & Newman, 2004; Newman & Evers, 2007). This particular finding demonstrates that, even though children with poorer language abilities may have some perceptual disadvantages (see introduction for more information about perceptual impairments), children with poorer language skills are nonetheless able to learn information about a talker's voice and transfer this knowledge to improving performance on a linguistic processing (here, spoken word recognition) task.

In addition to this positive result, the findings from the current study also indicate that children with poorer language skills are not able to generalize this knowledge of familiar talkers to other talkers with similar linguistic features (i.e., German-accented speech). Thus, although they may not be at a disadvantage when listening to familiar talkers, they are unable to generalize this knowledge to a novel talker. This could adversely affect processing and learning when listening to novel talkers in the classroom (e.g., substitute teachers) or when interacting with novel talkers in everyday life. It is possible that the inability to generalize to the speech of similar but novel talkers partially explains why children with language impairments have been found to have poorer speech perception ability (e.g., Shafer et al., 2005; Sussman, 1993, 2001) than children with typical language skills: Children with typically developing language map the novel talker in the experimental task to stored information about a similar talker to accurately complete the tasks, whereas children with language impairments do not. In terms of processing speech outside the laboratory, reduced generalization to novel talkers may mean that children with poorer skills are unable

to easily learn new language features (e.g., vocabulary, syntactic constructions) from unfamiliar talkers, which could, in turn, result in continued poorer language skills. In other words, difficulty in processing speech by novel talkers may explain, in part, the delayed development of language ability in these children.

In contrast, children with better language skills were able to learn something general about these foreign-accented talkers. This differs from earlier findings with a smaller set of participants that did not find any evidence for accent-general learning. The results found in the current study align with other studies showing that exposure to multiple talkers with the same foreign accent allows for generalization (Bradlow & Bent, 2008; Sidaras et al., 2009). The lack of generalization for the children with poorer language skills may be a result of needing additional practice and exposure. Previous research in children with poorer language skills found that these children required more exposure to generalize patterns and recognize regularities (Evans, Saffran, & Robe-Torres, 2009). Thus, it is possible that additional exposure and training would have allowed children with poorer language to generalize to novel yet similar accented talkers.

An additional important finding emerged from our study: A similar pattern of results was found for both high- and low-familiarity words. This is in contrast to the previous smaller study of Levi (2015) in which the FTA was only found for high-familiarity words. That same study examined an earlier data set with adults from Levi et al. (2011), which found no effect of lexical frequency on the FTA in adults. Thus, the current findings are more closely aligned to the adult data. Differences between the current study and our previous study may be due to the larger sample size and the inclusion of subject-dependent factors

such as language ability as predictors of performance. The inclusion of these predictors revealed the benefits of talker familiarity on the low-familiarity words. Thus, the current study suggests that talker familiarity should enhance novel word learning, as familiarity with a talker is associated with better perception of low-familiarity words in addition to high-familiarity words.

Together, these findings have broad implications for individuals with language impairments. If spoken language processing is facilitated for these listeners in certain contexts (e.g., with familiar talkers), it is possible that higher level linguistic processing would also be facilitated as a result of fewer cognitive resources going toward speech perception. Future research is necessary to determine whether talker familiarity in fact frees up cognitive resources and results in benefits to higher level linguistic processing.

Talker-Voice Learning, the FTA, and Accent-General Learning

Although not one of our objectives in conducting this study, our statistical analysis revealed an effect of talker voice learning. Previous research has found that only those listeners who learned the talkers sufficiently well (> 70% accuracy) demonstrated an FTA (Nygaard & Pisoni, 1998). Thus, exposure alone is not sufficient; listeners must actually demonstrate that they have learned information about a talker's voice (e.g., by identifying talkers 70% of the time) to reap the benefits in a spoken language processing task. The current study provides some support for this idea. Listeners in the current study learned the talkers very well by the final 2 days of training (interquartile range: 72.1%–86.9%, *Mdn* = 79.2%, full range: 40.0%–99.9%). Thus, nearly all listeners reached the previously set threshold of 70% accuracy, and all listeners, regardless of how well they learned the talkers, performed similarly on the word recognition task for familiar talkers.

The interesting finding to emerge from our analysis was that talker voice accuracy for the familiar talkers predicted accent-general transfer to the unfamiliar talkers. In other words, listeners who learned to identify the German-accented talkers well also performed well on the novel German-accented talkers, suggesting that they learned something general about German-accented speech. For these listeners, performance on the spoken word recognition task was similar for familiar and unfamiliar talkers. In contrast, those listeners who did not learn to identify the (familiar) talkers as well showed less accent-general transfer and did not perform as well for unfamiliar talkers on the spoken word recognition task.

Conclusion

The current study demonstrates that children across a range of language abilities are able to use stored information about a talker to improve spoken word recognition. Thus, the current study expands on previous studies of the FTA in listeners with typical language ability and

shows that even children with poorer language skills use stored information about familiar talkers to improve spoken language processing. Future research is necessary to determine whether talker familiarity frees up cognitive resources and results in benefits to higher level linguistic processing. Despite these positive findings, differences in perceptual skills did emerge; children with poorer language skills were less adept at generalizing to novel talkers who were similar to the trained, familiar talkers. Future research is needed to determine whether accent-general learning would occur with additional training. Similar results were found for both high- and low-familiarity words, suggesting that children with poorer language skills will benefit from learning novel words and possibly novel content in a school setting when information is presented by a familiar talker.

Acknowledgments

This work was supported by National Institute on Deafness and Other Communication Disorders Grant 1R03DC009851 to the first author. We would like to thank Gabrielle Alfano, Josh Barocas, Jennifer Bruno, Stephanie Lee, Emma Mack, Alexandra Muratore, Sydney Robert, and Margo Waltz for help with data collection and the children and their families for their participation.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Allen, J. S., & Miller, J. L. (2004). Listener sensitivity to individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America*, 115, 3171–3183.
- Barker, B. A., & Newman, R. S. (2004). Listen to your mother! The role of talker familiarity in infant streaming. *Cognition*, 94, B45–B53.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Benki, J. R. (2003). Quantitative evaluation of lexical status, word frequency, and neighborhood density as context effects in spoken word recognition. *The Journal of the Acoustical Society of America*, 113(3), 1689–1705.
- Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, 106(2), 707–729.
- Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *Test of Nonverbal Intelligence—Third Edition (TONI-3)*. Austin, TX: Pro-Ed.
- Case, J., Seyfarth, S., & Levi, S. V. (2018). Does implicit voice learning improve spoken language processing? Implications for clinical practice. *Journal of Speech, Language, and Hearing Research*, 61, 1251–1260. https://doi.org/10.1044/2018_JSLHR-L-17-0298
- Cortese, M. J., & Khanna, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods*, 40(3), 791–794.
- Dailey, N. S., Plante, E., & Vance, R. (2013). Talker discrimination in preschool children with and without specific language impairment. *Journal of Communication Disorders*, 46(4), 330–337.

- Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics*, 67, 224–238.
- Evans, J. L., Saffran, J. R., & Robe-Torres, K. (2009). Statistical learning in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*, 52, 321–335.
- Felty, R. A. (2007). *Context effects in spoken word recognition of English and German by native and non-native listeners* (Doctoral dissertation). University of Michigan, Ann Arbor, MI.
- Harel, D., & McAllister, T. (in press). Multilevel models for communication sciences and disorders. *Journal of Speech, Language, and Hearing Research*.
- Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, 56, 1–15.
- Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *The Journal of the Acoustical Society of America*, 29, 98–104.
- Levi, S. V. (2015). Talker familiarity and spoken word recognition in school-age children. *Journal of Child Language*, 42(4), 843–872.
- Levi, S. V., & Schwartz, R. G. (2013). The development of language-specific and language-independent talker processing. *Journal of Speech, Language, and Hearing Research*, 56, 913–920.
- Levi, S. V., Winters, S. J., & Pisoni, D. B. (2007). Speaker-independent factors affecting the perception of foreign accent in a second language. *The Journal of the Acoustical Society of America*, 121, 2327–2338.
- Levi, S. V., Winters, S. J., & Pisoni, D. B. (2011). Effects of cross-language voice training on speech perception: Whose familiar voices are more intelligible? *The Journal of the Acoustical Society of America*, 130(6), 4053–4062.
- Markham, D., & Hazan, V. (2004). The effect of talker- and listener-related factors on intelligibility for a real-word, open-set perception task. *Journal of Speech, Language, and Hearing Research*, 47, 725–737.
- McArthur, G. M., & Bishop, D. V. (2004). Which people with Specific Language Impairment have auditory processing deficits? *Cognitive Neuropsychology*, 21(1), 79–94.
- McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, 30(6), 1113–1126.
- McReynolds, L. V. (1966). Operant conditioning for investigating speech sound discrimination in aphasic children. *Journal of Speech and Hearing Disorders*, 9, 519–528.
- Newman, R. S., & Evers, S. (2007). The effect of talker familiarity on stream segregation. *Journal of Phonetics*, 35, 85–103.
- Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, 47, 204–238.
- Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, 60(3), 355–376.
- Nygaard, L. C., Sommers, M. S., & Pisoni, D. B. (1994). Speech perception as a talker-contingent process. *Psychological Science*, 5(1), 42–46.
- R Development Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>
- Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics*, 31, 509–527.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2007). E-Prime (Version 2.0 professional) [Computer software]. Pittsburgh, PA: Psychology Software Tools.
- Schroeder, M. R. (1968). Reference signal for signal quality studies. *The Journal of the Acoustical Society of America*, 44, 1735–1736.
- Semel, E., Wiig, E. H., & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals—Fourth Edition (CELF-4)*. Toronto, Ontario, Canada: The Psychological Corporation/Harcourt Assessment Company.
- Shafer, V. L., Morr, M. L., Datta, H., Kurtzberg, D., & Schwartz, R. G. (2005). Neurophysiological indexes of speech processing deficits in children with specific language impairment. *Journal of Cognitive Neuroscience*, 17(7), 1168–1180.
- Sidaras, S. K., Alexander, J. E., & Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *The Journal of the Acoustical Society of America*, 125(5), 3306–3316.
- Stark, R. E., & Heinz, J. M. (1996). Vowel perception in children with and without language impairment. *Journal of Speech and Hearing Research*, 39, 860–869.
- Sussman, J. E. (1993). Perception of formant transition cues to place of articulation in children with language impairments. *Journal of Speech and Hearing Research*, 36, 1286–1299.
- Sussman, J. E. (2001). Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions? *The Journal of the Acoustical Society of America*, 109(3), 1173–1180.
- Tallal, P., & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant perception. *Neuropsychologia*, 12, 83–93.
- Tallal, P., & Piercy, M. (1975). Developmental aphasia: The perception of brief vowels and extended stop consonants. *Neuropsychologia*, 13, 69–74.
- Tallal, P., & Stark, R. E. (1981). Speech acoustic-cue discrimination abilities of normally developing and language-impaired children. *The Journal of the Acoustical Society of America*, 69(2), 568–574.
- Theodore, R. M., & Miller, J. L. (2010). Characteristics of listener sensitivity to talker-specific phonetic detail. *The Journal of the Acoustical Society of America*, 128(4), 2090–2099.
- Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature*, 387, 176–178.
- Yonan, C. A., & Sommers, M. S. (2000). The effects of talker familiarity on spoken word identification in younger and older listeners. *Psychology and Aging*, 15(1), 88–99.

Appendix

Words Used in the Current Study

Words with high lexical familiarity.

bake	chop	hair	leash	peace	such
beach	coin	ham	leave	pet	tail
bean	comb	hang	lick	pill	then
bear	come	hatch	line	rash	these
beef	cop	hate	look	rat	thin
beer	couch	have	lose	ride	thumb
bib	date	head	mad	room	tight
big	dead	heel	map	rough	wag
bike	dish	here	match	run	web
bit	duck	him	meat	sale	whale
both	fair	hit	miss	sauce	what
bug	fat	hoop	moose	seed	where
buzz	fill	hot	mop	shake	white
cab	fish	house	mouse	shape	win
cake	fit	hug	neck	share	wing
cash	foot	jail	need	shave	wish
cat	fun	jar	net	should	with
cave	gas	kill	nice	shout	write
cheek	goose	kite	night	shut	wrong
cheese	gun	knife	page	son	year
choke	gym	leaf	pan	soon	

Words with low lexical familiarity.

bane	fate	hick	mime	rife	vole
bid	faze	hitch	mock	rile	wade
bile	feign	hock	mode	rut	wane
cad	fib	hone	mope	sage	wean
chafe	foul	jeer	muss	sake	whim
char	gab	knack	neap	sate	whip
chum	gape	lass	nick	sheath	whiz
cog	gauge	laud	node	shuck	wick
cope	gauze	leach	noose	siege	wit
core	gawk	ledge	notch	souse	womb
couth	gig	loathe	peal	sup	wraith
cuff	gin	loom	peer	tithe	wrath
cull	gnash	lore	pith	tone	wreath
dame	gnat	luff	poach	tout	zeal
debt	goad	lush	puck	vague	
dock	goon	mace	raid	vain	
doff	gull	maim	reap	vat	
dose	hail	mauve	reek	veil	
dung	haze	mesh	retch	void	