Generalized Learning of Dysarthric Speech between Male and Female Talkers

Micah E. Hirsch, M.S.[1], Kaitlin L. Lansford, Ph.D.[1], Tyson S. Barrett, Ph.D.,[2] Stephanie A. Borrie, Ph.D. [3]

[1]School of Communication Science and Disorders, Florida State University, Tallahassee, FL, USA

[2]Department of Psychology, Utah State University, Logan, UT, USA

[3]Department of Communicative Disorders and Deaf Education, Utah State University, Logan, UT, USA

Conflict of Interest Statement: The authors whose names are listed above certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Funding Statement: This research was supported by The American Speech-Language-Hearing Association's Students Preparing for Academic and Research Careers (SPARC) Award, awarded to Hirsch, and by the National Institute on Deafness and Other Communication Disorders Grant R21DC018867, awarded to Borrie (co-PI), Lansford (co-PI) and Barrett (co-I).

Correspondence: Micah E. Hirsch: mhirsch@fsu.edu

Abstract

Purpose: Perceptual training is a listener-targeted means for improving intelligibility of dysarthric speech. Recent work has shown that training with one talker generalizes to a novel talker of the same sex, and the magnitude of benefit is maximized when the talkers are perceptually similar. The current study expands previous findings by investigating whether perceptual training effects generalize between talkers of different sex.

Method: Forty new listeners were recruited for this study and completed a pretest, familiarization, and posttest perceptual training paradigm. Historical data collected using the same three-phase protocol were included in the data analysis. All listeners were exposed to the same talker with dysarthria during the pretest and posttest phases. For the familiarization phase, listeners were exposed to one of four talkers with dysarthria, differing in sex and level of perceptual similarity to the test talker, or a control talker. During the testing phases, listener transcribed phrases produced by the test talker with dysarthria. Listener transcriptions were then used to calculate a percent words correct intelligibility score.

Results: Multiple linear regression analysis revealed that intelligibility at posttest was not predicted by sex of the training talker. Consistent with earlier work, the magnitude of intelligibility gain was greater when the familiarization and test talkers were perceptually similar. Additional analyses revealed greater between listener variability in the dissimilar conditions as compared to the similar conditions.

Conclusions: Learning as a result of perceptual training with one talker with dysarthria generalized to another talker regardless of sex. In addition, listeners trained with perceptually similar talkers had greater and more consistent intelligibility improvement. Together, these

results add to previous evidence demonstrating that learning generalizes to novel talkers with dysarthria and that perceptual training is suitable for many listeners.

Generalized Learning of Dysarthric Speech between Male and Female Talkers

Behavioral treatment for dysarthria generally targets reduced intelligibility through speech production modifications. Common behavioral approaches, including cueing talkers with dysarthria to speak louder or clearer, have been shown to improve the listener's ability to understand the disordered speech (e.g. Hsu et al., 2019; Lam & Tjaden, 2016; Levy et al., 2017; Tjaden et al., 2013). Key, here, is that the burden of behavioral change is placed on the talker with dysarthria. Though previous research has demonstrated efficacy of talker-oriented management of dysarthria (e.g. Mahler & Ramig, 2012; Mahler et al., 2015; Ramig et al., 2001), not all individuals are able to behaviorally modify their production, and for others, such modifications do not adequately meet the talker's communication needs (Liss, 2007; Yorkston et al., 2017). To address this gap in clinical practice, additional or alternative treatment modalities to improve intelligibility need to be explored. Perceptual training offers a possible solution by shifting the weight of behavioral change from the talker onto the listener (Borrie et al, 2012a; Liss, 2007). Rather than requiring a talker to modify their speech to improve intelligibility, perceptual training exploits the malleability of the perceptual system to enhance a listener's ability to understand that speech signal via a familiarization experience.

Perceptual training is grounded in perceptual learning theory. The Ideal Adaptor Framework suggests that, during perceptual training, listeners learn distributional regularities present in multiple levels of acoustic information, including segmental and suprasegmental speech cues (Kleinschmidt & Jaeger, 2015). Based on these acoustic regularities, listeners form mental representations, or models, for different talkers that are updated with increased exposure, resulting in improved speech perception over time. Indeed, the preponderance of studies on perceptual training of dysarthric speech, to date, have revealed improved intelligibility following perceptual training (e.g., Borrie et al., 2012b; Borrie et al., 2017a; Kim, 2016; Kim & Nanney, 2014; Lansford et al., 2016; Lansford et al., 2018), suggesting that individualized training is a viable option for improving communication between a talker with dysarthria and their communication partners.

Listeners who interact with many individuals with dysarthria (e.g., healthcare workers), however, would benefit from perceptual training that generalizes to novel talkers. That is, improved perception of novel, untrained talkers with dysarthria following perceptual training with another talker. Such generalized learning would render talker-specific training for individuals who communicate with many different people with dysarthria unnecessary. Recent theoretical models of perceptual learning support such generalization; listeners leverage models formed from previous encounters with a specific talker to support their understanding of speech produced by a novel talker with shared perceptual characteristics (Kleinschmidt & Jaeger, 2015). Both top-down and bottom-up processes support the generalization process, such that listeners take advantage their model-driven expectations and the acoustic information present in the speech signal to facilitate their understanding of a novel talker. Thus, generalized learning occurs when there is likeness between the listeners' model-driven expectations and the incoming acoustic information.

Theoretical models of generalized learning are supported by empirical findings of improved processing of novel talkers following perceptual training with accented (Bradlow & Bent, 2008; Tzeng & Nygaard, 2012; Xie et al., 2020; Xie & Myers, 2017), degraded (Huyck et al., 2017) and, importantly, dysarthric speech (Borrie et al., 2017a). Perceptual adaptation to dysarthric speech not only generalizes to untrained talkers with dysarthria of the same sex, but is maximized when there is greater perceptual likeness between the training and test talkers, thereby supporting theoretical models (Borrie et al., 2017a). Indeed, many of the dysarthria subtypes share common speech features, such as slow speaking rate, imprecise consonants, compressed vowel space, and disordered vocal quality (Weismer & Kim, 2010). Thus, it is presumed that generalized adaption of dysarthric speech is broadly facilitated by such shared features, but is optimized when there is greater perceptual match between the training and test talkers. Although these findings are limited to male talkers with dysarthria, model assumptions suggest that learning should generalize across talkers, regardless of sex, if there is sufficient perceptual overlap between the listener's model-driven expectations and the incoming acoustic information (Kleinschmidt & Jaeger, 2015).

Although the Ideal Adaptor Framework posits that generalization between talkers of different sex should occur (Kleinschmidt & Jaeger, 2015), this is ultimately dependent on the speech cues listeners attend to during the learning process. General anatomical differences in the size and shape of vocal tracts associated with male and female talkers have well-documented impacts on the distribution of spectral information (Titze, 1989). Namely, fundamental and resonant frequencies in speech produced by male talkers tends to be lower than those of female talkers (Hillenbrand et al., 1995; Peterson & Barney, 1952). It is plausible, then, that generalizated learning between male and female talkers with dysarthria could be hindered by these anatomical and acoustic differences. However, a series of studies found that generalized learning of consonants produced by male and female talkers readily occurs when the distribution of speech cues was similar between the two speech samples, namely those that track to the temporal aspects of speech (Kraljic & Samuel, 2005, 2006, 2007). However, these studies were done with ambiguous speech samples at the phonemic level (i.e., training a novel phoneme), therefore generalization between talkers of different sex with dysarthria cannot be assumed.

The main research aim for this current study is to determine whether adaptation of dysarthric speech generalizes between talkers of different sex. The secondary aim is to determine whether perceptual similarity impacts generalized learning between male and female talkers with dysarthria. Specifically, the following research questions are addressed: 1) Does learning of dysarthric speech generalize to a novel talker of a different sex? 2) Will the effect of perceptual similarity, previously demonstrated in Borrie et al. (2017a), be robust to changes in the talker's sex? We hypothesized that familiarization effects will generalize between novel talkers of different sex evidenced by increased intelligibility scores on posttest. Also, and consistent with previous research, intelligibility gains are hypothesized to be greatest when the training and test talkers are perceptually similar to each other, regardless of talker sex.

Methods

Listeners

This study is based on data collected from 100 listener participants, including 40 newly recruited listeners and 60 listeners from a historical dataset (originally reported in Borrie et al., 2017a). All listeners used in this study, from both the newly recruited listeners and the Borrie et al. (2017a) dataset, were recruited in an indentical fashion via Amazon's Mechanical Turk (Amazon MTurk)[1]. Listeners were between 19 and 66 years old, reported no history or speech and language, or hearing disorders and were reimbursed $5 for their participation.

[1] Efficacy of the use of Amazon MTurk has been demonstrated by Lansford et al. (2016). Amazon MTurk settings used in this study limited participation those individuals with the "Master title" who had a 99% or higher HIT approval rating, have completed at least 500 approved HITs, and are located in the United States. Additionally, a qualification was added restricting individuals who already completed the study from participating again. Listeners recruited were also cross-referenced with listeners used in Borrie et al. (2017a) and other recent MTurk studies from our lab. Duplicate individuals were excluded from data analysis.

Talker Stimuli

Test Talker

During the pretest and posttest phases of this study, a total of 80 audio recorded phrases were presented to the listeners: 20 phrases in the pretest and 60 phrases in the posttest. The testing phrases ranged between three to five words in length. The testing phrases were comprised of real English words, but were semantically anomalous (e.g. *account for who could knock*, and *embark or take her sheet*) to prevent the use of top-down information to support perception of the stimuli. The test talker used in this study was an 84-year-old male with moderate ataxic dysarthria (see Table 1 for full description of the talker).

Familiarization Talkers

Audio recorded productions of an adapted version of the "Grandfather Passage" elicited by four talkers with dysarthria (see Table 1 for descriptions of the familiarization talkers) and one control talker were presented during the familiarization training phase of this study. The talkers with dysarthria were selected based on talker sex and level of similarity to the test talker. Level of similarity to the test talker was determined from results from Lansford et al. (2014)[2]. Listeners recruited for this study were assigned to either the female-similar or female-dissimilar group. Data for the male familiarization talkers and control conditions came from listeners who were recruited for the original study (Borrie et al., 2017a).

[2] In this study, listeners completed a free-classification task to group talkers together based on their perceptual features. Six clusters were formed from this task. The perceptually similar talkers used in this study came from the same cluster as the test talker, whereas the dissimilar talkers were chosen from a different cluster. Notable, these clusters were not formed based on medical etiology or the specific dysarthria diagnosis for a talker.

Table 1
Test and Familiarization Talker Profiles

Level of Similarity	**Age**	**Dysarthria subtype**	**Dysarthria Severity**	**Intel**	**Perceptual Characteristics***
Test talker	84	Ataxic	Moderate	49%	Reduced speaking rate, equal and even stress, prolonged phonemes and intervals, monotone, monoloudness, harsh vocal quality, and imprecise articulation with irregular articulatory breakdown
Female familiarization talkers (current study)					
Similar	63	Mixed	Moderate	43%	**Reduced speaking rate, equal and even stress, prolonged phonemes and intervals, monotone, monoloudness, harsh vocal quality and imprecise articulation**
Dissimilar	54	Hypokinetic	Moderate	60%	Normal speaking rate, short rushes of speech, reduced stress, **monotone, monoloudness, imprecise articulation**, hoarse vocal quality
Male familiarization talkers (Borrie et al., 2017)					
Similar	46	Mixed	Moderate	56%	**Reduced speaking rate, equal and even stress, prolonged phonemes and intervals, monotone, monoloudness, harsh vocal quality, and imprecise articulation**
Dissimilar	46	Ataxic	Moderate	47%	**Slightly reduced speaking rate**, reduced stress, **imprecise articulation**, **harsh vocal quality**
Control	67	-	-	-	

Notes: Talker information and intelligibility scores from Lansford and Liss (2014a, 2014b), and Lansford et al. (2014). *perceptual characteristics shared between the test and familiarization talkers are denoted in bold font. Intel = Intelligibility

Procedure

From Amazon MTurk, listeners were directed to a secured web-browser hosted by Utah State University to complete the study. Prior to starting the experiment, listeners were instructed to wear earphones during the experiment. After obtaining informed consent and completing a brief demographic questionnaire, all listeners recruited for this current study completed the three-phase experimental procedure: pretest, familiarization, and posttest.

During the pretest phase, listeners were presented with the audio phrase recordings, one at a time, and instructed to type out what they thought they heard. They were encouraged to try and guess, if they were not sure, and to type in an “X” for any unintelligible utterance. The pretest phase was self-paced. Listeners had to press “Next” in order to hear and transcribe the next phrase. Phrases were only presented once. Following the pretest, listeners engaged in the familiarization phase, in which they were instructed to listen to either the female similar or female dissimilar talker’s reading of the “Grandfather Passage” while reading along to the text. The “Grandfather Passage” was presented in 35 phrases. Similar to the pretest, listeners had to press “Next” in order to hear the next phrase from the passage. Following the familiarization phase, the listeners completed the posttest. The instructions and procedures for completing the post-test were the same as the pretest. As with the pretest, the posttest was self-paced.

Data collected from the original study (Borrie et al., 2017a) were also analyzed along with the new data collected in this study. The procedures used in the Borrie et al. (2017a) study were the same to the procedures described above, with the exception that the listeners for the original study were exposed to speech produced by either a healthy control, dissimilar male, or similar male talker during the familiarization phase. Other than the different familiarization talkers, the procedures, the types of stimuli presented to the listeners, and the web-based application were identical to those used in this current study.

Data Analysis

The number of correct words transcribed during the pretest and posttest were counted using an automated, open-sourced scoring program, Autoscore[3] (www.autoscore.usu.edu [Borrie et al., 2019]). A word was scored as correct if it matched the intended target word exactly or only differed by a tense or plural marking. Misspellings, homophones, and substitutions between the words *a* and *the* were also scored as correct. A pretest and posttest percent words correct (PWC) score was then generated by dividing the number of correct words by the total number of words possible and multiplying by 100. Next, the difference between the PWC scores in the pre-test and post-test phases was calculated to obtain an improvement PWC score for each listener.

To assess whether there were differences in PWC across the experimental conditions, we used linear regression with the posttest PWC score predicted by each familiarization talker controlling for pretest PWC score. This specification results in coefficients that are in reference to differences in improvement from pretest to posttest. The control condition was set as the reference in the regression. In order to fully compare differences in improvement in PWC by the talker sexes and the similarity levels, we used linear contrasts. These comparisons also highlighted which factors resulted in more consistent improvement across listeners. Assumptions of normality and homoscedasticity were evaluated to ensure linear regression was an appropriate method for data analysis. Additionally, a coefficient of variation test was conducted to examine the variation differences in the magnitude of intelligibility improvement associated with the talker sex and similarity manipulations.

[3] Autoscore has been validated as an accurate (99% accuracy) and efficient scoring tool (Borrie et al., 2019). Thus reliability measures for scoring the transcripts in this study was not deemed necessary.

Results

The estimates from the multiple regression analysis, highlighting the differences between the experimental conditions and the control condition, controlling for pretest, are shown in Table 2. All experimental conditions were significantly different from the control group when controlling for pretest scores. Thus, the listeners who were familiarized with a talker with dysarthria, on average, had greater intelligibility improvement compared to listeners who were trained with the control talker (see Figure 1). The full model accounted for approximately 55% of the variance in the outcome. Talker sex and similarity level alone accounted for 33% of the variance in the outcome. As such, a large portion of the intelligibility scores on posttest can be explained by the familiarization conditions. Notably, with pretest included as a covariate, neither normality nor homoscedasticity were violated and indicate multiple linear regression was appropriate for analysis.

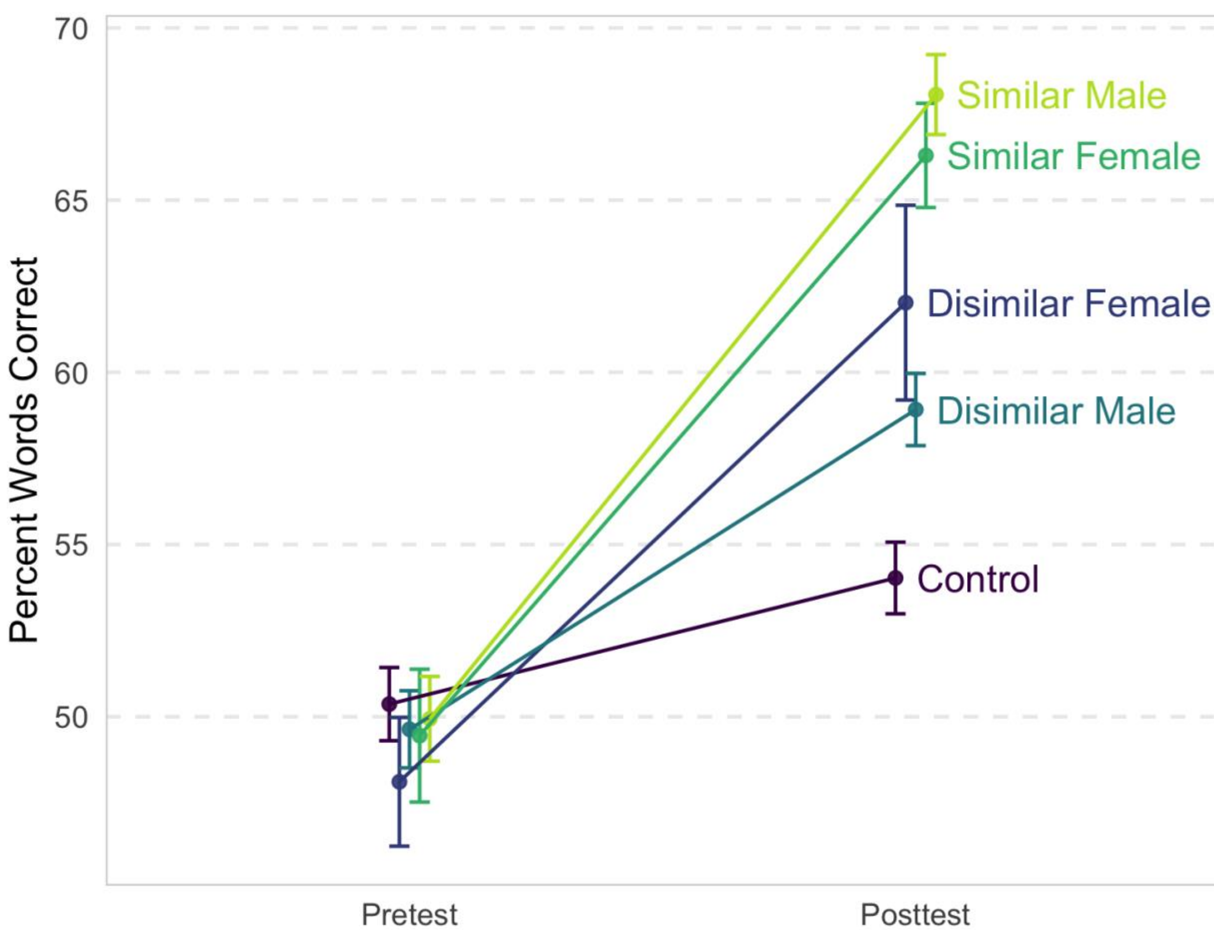
Table 2
Regression Model Results for Generalized Adaptation

	PWC	SE	p-value
Control	ref	ref	ref
Similar Male	14.310	1.9306	< .001
Similar Female	12.859	1.9321	< .001
Dissimilar Male	5.3613	1.9314	.0066
Dissimilar Female	9.445	1.9417	< .001
Pretest	0.6432	0.0939	< .001
R^2 Full Model	0.552		
R^2 Without Pre-Test	0.328		

Figure 1

Pretest and Posttest Percent Words Correct Scores by Talker Condition

Note: The mean pretest and posttest percent words correct (PWC) scores for each condition is depicted above. Data from the control, male similar, and male dissimal conditions came from the original study (Borrie et al., 2017a). Data for the female similar and dissimilar conditions were collected for this current study. Notably, all listeners trained with dysarthric talkers had greater PWC scores on posttest compared to listeners trained with a control talker. Listeners trained with similar talkers had the greatest PWC improvement compared to the listeners trained with dissimilar talkers.



Talker Sex

To examine the whether there were significant differences between the improvement in PWC based on talker sex, linear contrasts were used to compare male and female training talker conditions. Overall, improvement in PWC was not impacted by the sex of the familiarization talker, such that training with a male and female talker yielded comparable results (p = .34). Additional testing, factoring in the effect of similarity, revealed there was no significant difference between the male and female similar conditions (p = 0.4541). However, a significant difference was found between the male and female dissimilar conditions (p = 0.0375), with greater intelligibility improvement revealed for the female dissimilar condition. The coefficient of variation analysis revealed that the variation in intelligibility improvement (change in intelligibility from pretest to posttest) was equivalent for listeners familiarized with male or female talkers (see Table 3).

Table 3
Comparisons between Variation in Learning

	Between Level of Similarity		Between Sex		Between Conditions	
Test	D'AD	p-value	D'AD	p-value	D'AD	p-value
Asymptotic	22.73	< .001*	3.83	.15	20.78	< .001*
MSLR	24.96	< .001*	2.98	.23	26.45	< .001*

Note. Significant values are marked with an asterisk.

Similarity

Linear contrasts were conducted to examine the effect of similarity of the training and test talkers on improvement in PWC. Overall, a significant difference in improvement was found for listeners familiarized with similar talkers vs. dissimilar talkers, such that training with perceptually similar talkers resulted in higher improvement in intelligibility (p < .001). Additional testing, accounting for talker sex, revealed a significant difference between the male similar and dissimilar conditions ($p < .001$; similar > dissimilar). Furthermore, the difference between the female similar and dissimilar conditions were significant at the .10 alpha level ($p = .08$; similar > dissimilar).

The distribution of listener improvement scores in PWC based on the level of similarity, shown in Figure 2, shows that the listeners familiarized with dissimilar talkers had higher variability in change from pretest to posttest, as compared to both the control and similar conditions. The results for the coefficient of variation asymptotic and MSLR tests are reported in Table 3. These results indicate that familiarization with the similar talkers resulted in more consistent intelligibility improvement than familiarization with the dissimilar talkers.

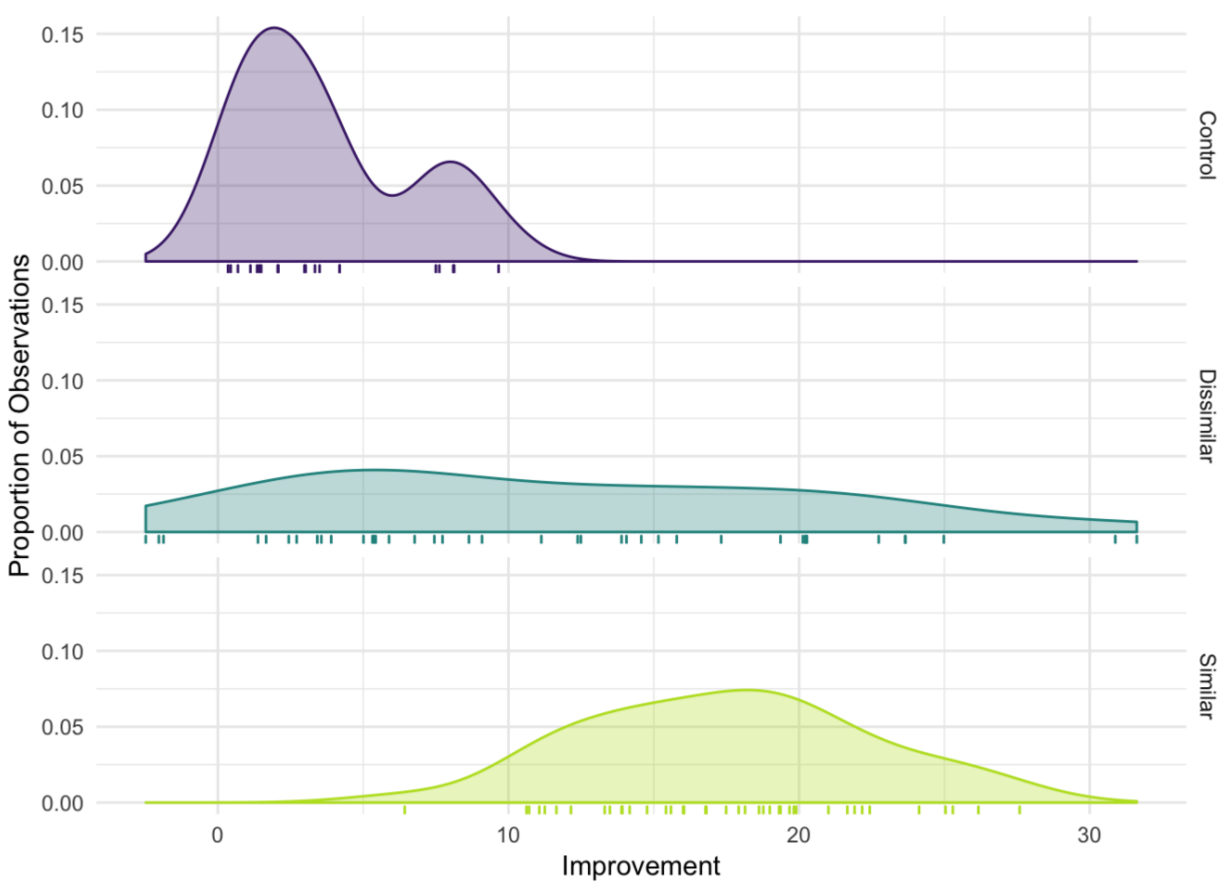
Discussion

The primary goal of the current study was to investigate whether perceptual learning of dysarthric speech (Borrie et al., 2017a) generalizes between talkers of different sex. Overall, the results from the current study suggest this is the case; listeners trained with a female talker with dysarthria improved their ability to understand a male talker with dysarthria. In addition, and consistent with the original study by Borrie and colleagues (2017a), the current work revealed that generalized learning of dysarthric speech was constrained by level of perceptual similarity between the training and test talker; intelligibility improvement of the test talker was greatest for listeners trained with a perceptually similar talker, irrespective of the training talker's sex.

Figure 2

Distribution of Improvement Scores Paneled by Level of Similarity

Note: The above figure shows the intelligibility improvement scores for listeners from both the original study (Borrie et al., 2017a) and this current investigation. The figures show distributions of intelligibility improvement across talker sex, with improvement scores from listeners trained with the control talker (top), the dissimilar talkers (middle), and the similar talkers (bottom). The listeners trained with the dissimilar talkers showed greater variability in intelligibility improvement compared to the listeners trained with the similar talkers, evidenced by a flatter distribution and by coefficient of variation analysis results (Table 3).



Results from this study are consistent with current models of speech perception and perceptual learning. Recall, the Ideal Adaptor Framework posits listeners leverage generative models formed through experience with other talkers to process the speech of novel talkers. Generalized learning occurs via a decision-making process utilizing both model-driven expectations and the acoustic information presented in the speech to determine whether generalization of the model is appropriate. Thus, generalization should occur between talkers of different sex if there is sufficient shared structure between the two speech signals (Kleinschmidt & Jaeger, 2015). The current results suggest that most listeners were able to apply generative models formed from both male and female talkers with dysarthria to a novel talker, and that the magnitude of intelligibility improvement was mediated by the level of perceptual similarity between the training and test talkers.

The current results demonstrate that listeners are able to detect shared structure in the acoustic cues between male and female talkers with dysarthria, and that this effect was magnified for talkers who were perceptually similar to each other. Previous research has found that generalization of perceptual learning effects is dependent on the level of shared structure between talkers of different sex. Kraljic and Samuel found perceptual learning effects at the phonemic level generalized between talkers of different sex for stop consonants (2006, 2007), but not fricatives (2005, 2007). The authors speculated that their findings were likely due to the distribution of temporal and spectral cues offered by the different consonants. Namely, stop consonants produced by male and female talkers share similar voice onset time (a temporal cue), while fricative consonants differ in their spectral characteristics (Kraljic & Samuel, 2007). Although this current study does not investigate the acoustic similarities between the talkers, Lansford et al. (2014) found that the perceptual similarity ratings used by the current study to define the similar and dissimilar categories were strongly correlated with acoustic measures that track to temporal aspects of speech (e.g., speaking rate and rhythm). Thus, it is possible that the generalized adaptation observed in our series of studies could be driven by overlapping distributions of acoustic information in the temporal domain. Future research in this area should systematically examine the role of acoustics in generalized adaptation to dysarthric speech.

Interestingly, training with talkers who were more perceptually similar to the test talker not only led to greater perceptual gains overall, but also led to a less variable learning effect across listeners, as evidenced by the coefficient of variation analysis. As compared to the stable magnitude of intelligibility improvement measured in the similar conditions, the magnitude of improvement from pretest to posttest following training with the dissimilar speakers varied widely across listeners, with many listeners experiencing little to no perceptual benefit. These results suggest that some listeners may be better equipped than others to form and utilize generative models when there is less distributional overlap between the training and test talkers. Certainly, previous research has supported that generalized adaptation of speech relies on the amount of similarity between speech samples (e.g. Kleinschmidt & Jaeger, 2015; Xie et al., 2020). However, it is unknown why some listeners are better able to recognize similar structures across talkers to support perception of novel, yet perceptually dissimilar, talkers. It is likely, though, that generalization effects are impacted by additional factors not measured here. Indeed, recent evidence suggests that listener-related factors, such as age, hearing status, rhythm perception abilities, and other cognitive processes impact perception of and adaptation to dysarthric speech (Borrie et al. , 2017b, 2018; Ingvalson et al., 2017a; Lansford et al., 2019; McAuliffe et al., 2013). Further, research has not yet systematically considered how adaptation and generalization are impacted by listener attitudes and level of comfort with disordered speech, despite evidence suggesting that these factors may be related (Guo & Togher, 2008; Ingvalson et al., 2017b). It is plausible, then, that the ability to form generative models during the familiarization process and apply them to novel talkers may be differentially impacted by such listener-related factors, and, as such, should be accounted for in future work in this area.

In the current study, a significant difference was found between the dissimilar male and dissimilar female conditions. Listeners trained with the dissimilar female talker demonstrated greater intelligibility improvement at posttest compared to those listeners familiarized with the dissimilar male talker. This is an interesting, and unanticipated, finding that may be explained by the difference in overall intelligibility levels of the familiarization talkers selected for this study. The dissimilar female talker's overall intelligibility level was 60%, while the male dissimilar talker's was 47% (originally reported in Lansford & Liss, 2014a). As revealed by Borrie et al. (2017a), in addition to level of perceptual similarity between the test and familiarization talkers, overall intelligibility level of the familiarization talker also constrained the magnitude of intelligibility improvement following familiarization, such that listeners who were trained with talkers with mid to high levels of intelligibility demonstrated greater perceptual processing of the novel talker than those trained with talkers with low levels of intelligibility (Borrie et al., 2017a). Thus, it is possible that, in the current study, overall intelligibility of the familiarization talkers may partially explain the difference in performance between the listeners in the male dissimilar and female dissimilar training groups. Additionally, level of similarity between two speakers is unlikely to be a binary variable; it is more likely that dissimilarity and similarity is continuous. Therefore, dichotomous categorization of similarity may have also impacted the results of this study. Research should continue to investigate how the level of intelligibility and perceptual match between familiarization and test talkers' mediate familiarization of dysarthric speech.

Results from this study also continue to demonstrate the efficacy of perceptual training as a potential adjunct to speech treatment for talkers with dysarthria. Current evidence shows promise for the use of perceptual training programs for any person who interacts with an individual, or individuals, with dysarthria. Specifically, the generalized effects from familiarization can lead to the development of training programs for individuals who interact with multiple talkers with dysarthria, such as healthcare workers. However, additional research is required in order to evaluate the effectiveness of familiarization training paradigms implemented in clinical settings. Thus, future studies should investigate perceptual training effects, and the generalization of such effects, in clinical settings with listeners who will potentially interact with talkers with dysarthria, such as physicians, physical therapists, and occupational therapists.

Conclusion

Overall, this study extended results regarding generalized adaptation of dysarthric speech by showing that learning generalizes to talkers of a different sex. Further, the results of this study confirmed previous results regarding the key role that perceptual similarity plays in generalized learning. Taken together, the findings from this study have potential clinical implications for the development of generalized perceptual training programs that could be implemented along with traditional speech treatment for improved management of the intelligibility deficits that characterize dysarthria.

Acknowledgements

This research was done as part of the first author’s master’s thesis at Florida State University and was funded by The American Speech-Language-Hearing Association’s Students Preparing for Academic and Research Careers (SPARC) Award, awarded to Hirsch, and by the National Institute on Deafness and Other Communication Disorders Grant R21DC018867, awarded to

Borrie (co-PI), Lansford (co-PI) and Barrett (co-I). We also extend our gratitude to Dr. Julie Liss at Arizona State University for continued use of her extensive dysarthria speech sample database.

References

Borrie, S. A., Barrett, T. S., & Yoho, S. E. (2019). Autoscore: An open-source automated tool for scoring listener perception of speech. *The Journal of the Acoustical Society of America*, *145*(1), 392–399. https://doi.org/10.1121/1.5087276

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2017a). Generalized Adaptation to Dysarthric Speech. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3110–3117. https://doi.org/10.1044/2017_jslhr-s-17-0127

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2017b). Rhythm Perception and Its Role in Perception and Learning of Dysrhythmic Speech. *Journal of Speech, Language, and Hearing Research*, *60*(3), 561–570. https://doi.org/10.1044/2016_jslhr-s-16-0094

Borrie, S. A., Lansford, K. L., & Barrett, T. S. (2018). Understanding dysrhythmic speech: When rhythm does not matter and learning does not happen. *The Journal of the Acoustical Society of America*, *143*(5), EL379–EL385. https://doi.org/10.1121/1.5037620

Borrie, S. A., McAuliffe, M. J., & Liss, J. M. (2012). Perceptual Learning of Dysarthric Speech: A Review of Experimental Studies. *Journal of Speech, Language, and Hearing Research*, *55*(1), 290–305. https://doi.org/10.1044/1092-4388(2011/10-0349)

Borrie, S. A., McAuliffe, M. J., Liss, J. M., Kirk, C., O'Beirne, G. A., & Anderson, T. (2012). Familiarisation conditions and the mechanisms that underlie improved recognition of dysarthric speech. *Language and Cognitive Processes*, *27*(7–8), 1039–1055. https://doi.org/10.1080/01690965.2011.610596

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729. https://doi.org/10.1016/j.cognition.2007.04.005

Guo, Y. E., & Togher, L. (2008). The impact of dysarthria on everyday communication after traumatic brain injury: A pilot study. *Brain Injury*, *22*(1), 83–97. https://doi.org/10.1080/02699050701824150

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, *97*(5), 3099–3111. https://doi.org/10.1121/1.411872

Hsu, S. C., McAuliffe, M. J., Lin, P., Wu, R.-M., & Levy, E. S. (2019). Acoustic and Perceptual Consequences of Speech Cues for Mandarin Speakers With Parkinson's Disease. *American Journal of Speech-Language Pathology*, *28*(2), 521–535. https://doi.org/10.1044/2018_AJSLP-18-0020

Huyck, J. J., Smith, R. H., Hawkins, S., & Johnsrude, I. S. (2017). Generalization of Perceptual Learning of Degraded Speech Across Talkers. *Journal of Speech, Language, and Hearing Research*, *60*(11), 3334–3341. https://doi.org/10.1044/2017_jslhr-h-16-0300

Ingvalson, E. M., Lansford, K. L., Federova, V., & Fernandez, G. (2017). Listeners' attitudes toward accented talkers uniquely predicts accented speech perception. *The Journal of the Acoustical Society of America*, *141*(3), EL234–EL238. https://doi.org/10.1121/1.4977583

Ingvalson, E. M., Lansford, K. L., Fedorova, V., & Fernandez, G. (2017). Receptive vocabulary, cognitive flexibility, and inhibitory control differentially predict older and younger adults' success perceiving speech by talkers with dysarthria. *Journal of Speech Language and Hearing Research*, *60*(12), 3632–3641. https://doi.org/10.1044/2017_JSLHR-H-17-0119

Kim, H. (2016). Familiarization effects on consonant intelligibility in dysarthric speech. *Folia Phoniatrica et Logopaedica*, *67*(5), 245–252. https://doi.org/10.1159/000444255

Kim, H., & Nanney, S. (2014). Familiarization effects on word intelligibility in dysarthric speech. *Folia Phoniatrica et Logopaedica*, *66*(6), 258–264.

https://doi.org/10.1159/000369799

Kleinschmidt, D. F., & Florian Jaeger, T. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203. https://doi.org/10.1037/a0038695

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*(2), 141–178. https://doi.org/10.1016/j.cogpsych.2005.05.001

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin and Review*, *13*(2), 262–268. https://doi.org/10.3758/BF03193841

Kraljic, T., & Samuel, A. G. (2007). Perceptual adjustments to multiple speakers. *Journal of Memory and Language*, *56*(1), 1–15. https://doi.org/10.1016/j.jml.2006.07.010

Lam, J., & Tjaden, K. (2016). Clear Speech Variants: An Acoustic Study in Parkinson's Disease. *Journal of Speech, Language, and Hearing Research*, *59*(4), 631–646. https://doi.org/10.1044/2015_jslhr-s-15-0216

Lansford, K. L., Borrie, S. A., & Barrett, T. S. (2019). Regularity matters: Unpredictable speech degradation inhibits adaptation to dysarthric speech. *Journal of Speech Language and Hearing Research*, *62*(12), 4282–4290. https://doi.org/10.1044/2019_JSLHR-19-00055

Lansford, K. L., Borrie, S. A., & Bystricky, L. (2016). Use of crowdsourcing to assess the ecological validity of perceptual-training paradigms in dysarthria. *American Journal of Speech-Language Pathology*, 1–7. https://doi.org/10.1044/2015

Lansford, K. L., & Liss, J. M. (2014). Vowel acoustics in dysarthria: Mapping to perception. *Journal of Speech, Language, and Hearing Research*, *57*, 68–80.

Lansford, K. L., Luhrsen, S., Ingvalson, E. M., & Borrie, S. A. (2018). Effects of familiarization on intelligibility of dysarthric speech in older adults with and without hearing loss.

American Journal of Speech-Language Pathology, *27*(1), 91–98. https://doi.org/10.1044/2017_AJSLP-17-0090

Levy, E. S., Chang, Y. M., Ancelle, J. A., & McAuliffe, M. J. (2017). Acoustic and perceptual consequences of speech cues for children with dysarthria. *Journal of Speech Language and Hearing Research*, *60*(6S), 1766–1779. https://doi.org/10.1044/2017_JSLHR-S-16-0274

Liss, J. M. (2007). The role of speech perception in motor speech disorders. In G. Weismer (Ed.), *Motor Speech Disorders: Essays for Ray Kent* (pp. 186–219). Plural Publishing Inc.

Mahler, L. A., & Ramig, L. O. (2012). Intensive treatment of dysarthria secondary to stroke. *Clinical Linguistics and Phonetics*, *26*(8), 681–694. https://doi.org/10.3109/02699206.2012.696173

Mahler, L. A., Ramig, L. O., & Fox, C. (2015). Evidence-based treatment of voice and speech disorders in Parkinson disease. *Current Opinion in Otolaryngology and Head and Neck Surgery*, *23*(3), 209–215. https://doi.org/10.1097/MOO.0000000000000151

Mcauliffe, M. J., Gibson, E. M. R., Kerr, S. E., Anderson, T., & Lashell, P. J. (2013). Vocabulary influences older and younger listeners' processing of dysarthric speech. *The Journal of the Acoustical Society of America*, *134*(2), 1358–1368. https://doi.org/10.1121/1.4812764

Peterson, G. E., & Barney, H. L. (1952). Control Methods Used in a Study of the Vowels. *Journal of the Acoustical Society of America*, *24*(2), 175–184. https://doi.org/10.1121/1.1906875

Ramig, L. O., Sapir, S., Countryman, S., Pawlas, A. A., O'brien, C., Hoehn, M., & Thompson, L. L. (2001). Intensive voice treatment (LSVT®) for patients with Parkinson's disease: a 2 year follow up. *Journal of Neurology, Neurosurgery, and Psychiatry*, *71*, 493–498.

https://doi.org/10.1136/jnnp.71.4.493

Titze, I. R. (1989). Physiologic And Acoustic Differences Between Male And Female Voices. *Journal of the Acoustical Society of America*, *85*(4), 1699–1707. https://doi.org/10.1121/1.397959

Tjaden, K., Richards, E., Kuo, C., Wilding, G., & Sussman, J. (2013). Acoustic and perceptual consequences of clear and loud speech. *Folia Phoniatrica et Logopaedica*, *65*(4), 214–220. https://doi.org/10.1159/000355867

Tzeng, C. Y., & Nygaard, L. C. (2012). The effect of training structure on perceptual learning of accented speech. *The Journal of the Acoustical Society of America*, *131*(4), 3310–3310. https://doi.org/10.1121/1.4708383

Weismer, G., & Kim, Y. (2010). Classification and taxonomy of motor speech disorders: What are the issues? In B. Maassen & P. van Lieshout (Eds.), *Speech Motor Control: New Developments in Basic and Applied Research* (pp. 229–242). OUP Oxford.

Xie, X., Liu, L., & Jaeger, T. F. (2020). Cross-talker generalization in the perception of non-native speech : a large-scale replication. *Open Science Framework*, 1–79. https://doi.org/10.17605/OSF.IO/BRWX5

Xie, X., & Myers, E. B. (2017). Learning a talker or learning an accent: Acoustic similarity constrains generalization of foreign accent adaptation to new talkers. *Journal of Memory and Language*, *97*, 30–46. https://doi.org/10.1016/j.jml.2017.07.005

Yorkston, K., Baylor, C., & Britton, D. (2017). Speech versus speaking: The experiences of people with parkinson’s disease and implications for intervention. *American Journal of Speech-Language Pathology*, *26*(2Special Issue), 561–568. https://doi.org/10.1044/2017_AJSLP-16-0087