JSLHR

Research Article

Generalized Adaptation to Dysarthric Speech

Stephanie A. Borrie,[a] Kaitlin L. Lansford,[b] and Tyson S. Barrett[c]

Purpose: Generalization of perceptual learning has received limited attention in listener adaptation studies with dysarthric speech. This study investigated whether adaptation to a talker with dysarthria could be predicted by the nature of the listener's prior familiarization experience, specifically similarity of perceptual features, and level of intelligibility.
Method: Following an intelligibility pretest involving a talker with ataxic dysarthria, 160 listeners were familiarized with 1 of 7 talkers with dysarthria—who differed from the test talker in terms of perceptual similarity (same, similar, dissimilar) and level of intelligibility (low, mid, high)—or a talker with no neurological impairment (control). Listeners then completed an intelligibility posttest on the test talker.
Results: All listeners benefited from familiarization with a talker with dysarthria; however, adaptation to the test talker was superior when the familiarization talker had similar perceptual features and reduced when the familiarization talker had low intelligibility.
Conclusion: Evidence for both generalization and specificity of learning highlights the differential value of listeners' prior experiences for adaptation to, and improved understanding of, a talker with dysarthria. These findings broaden our theoretical knowledge of adaptation to degraded speech, as well as the clinical application of training paradigms that exploit perceptual processes for therapeutic gain.

Listener-targeted perceptual learning paradigms can reduce the intelligibility burden of dysarthria. In a series of empirical studies, we have demonstrated that prior experience, or *familiarization*, with dysarthric speech yields improved understanding of the same talker(s) in subsequent encounters (e.g., Borrie et al., 2012; Borrie, McAuliffe, Liss, O'Beirne, & Anderson, 2013; Borrie & Schäfer, 2015, 2017; Lansford, Borrie, & Bystricky, 2016). Successful speech perception requires that listeners map the incoming auditory input onto stored linguistic categories, and while the precise learning mechanisms underlying the adaptive benefits of familiarization have not been fully realized, it is hypothesized that experience with a talker with dysarthria affords listeners an opportunity to retune their stored categories by mapping representations of the noncanonical, or degraded, acoustic cues (Liss, Spitzer, Caviness, & Adler, 2002). Evidence that the speech perception system can rapidly and flexibly adapt to noncanonical representations of speech has also been observed with foreign accents (e.g., Clarke & Garrett, 2004) and artificially degraded signals including synthetic (e.g., Francis, Ciocca, Ma, & Fenn, 2008), noise-vocoded (e.g., Davis, Johnsrude, Hervais-Adelman, Taylor, & McGettigan, 2005), and time-compressed speech (e.g., Golomb, Peelle, & Wingfield, 2007).

According to Kleinschmidt and Jaeger's (2015) ideal adaptor framework, perceptual learning can be conceptualized as a form of statistical inference: listeners build linguistic generative models for novel talkers based on knowledge about the distribution of acoustic cues associated with each linguistic category. Given the probabilistic nature of cue-to-category mapping, processing the signal in subsequent encounters with the same talker is then considered to be a process of "inference under uncertainty" (p. 150). However, the framework also proposes that the speech perception system is sensitive to structure over talkers and similar situations, and can draw on previous experience to support current processing. In this sense, the generative model and distributional beliefs developed and updated during encounters with one talker may generalize to improved understanding of a novel talker with a similar way of speaking.

A primary goal of listener-targeted remediation in dysarthria is to offset the communicative burden from the talker on to their primary communication partner (e.g.,

[a]Department of Communicative Disorders and Deaf Education, Utah State University, Logan
[b]School of Communication Sciences and Disorders, Florida State University, Tallahassee
[c]Department of Psychology, Utah State University, Logan
Correspondence to Stephanie A. Borrie: stephanie.borrie@usu.edu

Editor-in-Chief: Krista Wilkinson
Editor: Jeannette Hoit
Received April 9, 2017
Revision received May 30, 2017
Accepted June 2, 2017
https://doi.org/10.1044/2017_JSLHR-S-17-0127

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

caregiver, spouse, friend; Borrie, McAuliffe, & Liss, 2012). Thus, studies in the area of perceptual learning of dysarthric speech have focused on talker-specific adaptation, whereby listeners are familiarized and tested on the same talker or the same small group of talkers with relatively homogenous patterns of degradation. However, health care professionals who regularly interact with patients with dysarthria (e.g., physicians, nurses, and physical therapists) could also benefit from listener-targeted remediation, particularly if learning is generalized to novel talkers. To date, only one study affords preliminary evidence of generalized, or talker-independent, adaptation to dysarthria speech (Liss et al., 2002). In the study by Liss et al. (2002), listeners were familiarized with a group of talkers exhibiting either hypokinetic or ataxic dysarthria. Two posttests then ensued: the first test consisted of testing phrases produced by the same-talker group encountered during familiarization and the second test consisted of a small subset of the testing phrases produced by the novel talker group (i.e., hypokinetic dysarthria if familiarized with ataxic dysarthria). The posttest intelligibility data revealed that familiarization facilitated adaptation to all talkers with dysarthria; however, performance was superior for the talkers encountered during familiarization. It is, however, important to acknowledge the limitation of phrase duplication and intelligibility discrepancies across the two test conditions.[1] Evidence of generalized adaptation has also been reported with other forms of noncanonical speech, including accented and time-compressed signals (e.g., Baese-Berk, Bradlow, & Wright, 2013; Bradlow & Bent, 2008; Pallier, Sebastián-Gallés, Dupoux, Christophe, & Mehler, 1998). Taken together, we speculate that listeners are sensitive to shared structure over talkers with dysarthria and can benefit from statistical knowledge of such.

In a comprehensive review of the literature, Weismer and Kim (2010) identified a subset of perceptual features that unify the dysarthrias as a group of neurogenic speech disorders. These general dysarthria features, common in the speech patterns of most talkers with dysarthria, include reduced vowel space, reduced speaking rate, reduced formant transitions and articulatory speed, and reduced phonetic contrasts. Of course, the literature also documents dysarthria-specific perceptual features, widely considered to be tied to a talker's underlying medical etiology (Darley, Aronson, & Brown, 1975; Duffy, 2013). Although, recent work has initiated efforts to identify dysarthria-specific features through the study of perceptual similarity rather than underlying medical etiology (Lansford, Liss, & Norton, 2014; Lansford, Berisha, & Utianski, 2016). We postulate that listeners can exploit distributional knowledge about perceptual features, learned during experience with one talker with dysarthria, for improved processing of another. Further, we predict that the generative model established during prior experience will be most advantageous when applied to a novel talker with a similar-sounding dysarthria. In such situations, listeners can take advantage of structured regularities in both dysarthria-general and dysarthria-specific features. A systematic investigation into the nature of prior experiences for adaptation to, and improved understanding of, dysarthric speech is required to advance this theory.

Talkers with dysarthria present with different levels of intelligibility, defined as the success with which the talker's message is recovered by a listener (Hustad, 2008; Yorkston, Beukelman, Strand, & Bell, 1999). Guediche, Fiez, and Holt (2016) have noted the increased challenge of mapping severely degraded acoustic cues to stored linguistic categories, advancing the idea that alignment of cues for perceptual learning of speech is aided by a more intelligible signal. Thus, speech signals with low levels of intelligibility may disrupt cue-to-category mapping, compromising learning of the degraded acoustic cues and, consequently, the ability to integrate prior knowledge with subsequent processing. Therefore, an exploration of whether the talker's level of intelligibility during familiarization affects adaptation is also warranted.

The purpose of this study was to investigate whether prior experience with a talker with dysarthria generalizes to improved understanding of another talker with dysarthria by systematically modifying the listener's experience (i.e., familiarization talker) in terms of similarity of perceptual features (same, similar, dissimilar) and level of intelligibility (low, mid, high). Specifically, the following two key research questions were addressed: (a) Does similarity of perceptual features during familiarization predict adaptation to a talker with dysarthria? (b) Does level of intelligibility during familiarization predict adaptation to a talker with dysarthria? We hypothesized that adaptation to a talker with dysarthria, quantified as intelligibility performance following familiarization, would be differentially influenced by the nature of the listener's prior experience—that adaptation would be superior when the familiarization talker had similar perceptual features and reduced when the familiarization talker had low intelligibility.

Method

Listener Participants

One hundred and sixty young adults (83 women, 77 men) aged 19 to 45 years old (M = 31.04, SD = 5.29) participated in the experiment. All participants were native speakers of American English and currently residing in the United States. Participants reported no history of speech, language, or hearing problems and no significant prior contact with persons having motor speech disorders.

Participants were recruited using the crowdsourcing website Amazon Mechanical Turk (MTurk; http://www.

[1]Due to the small set of experimental stimuli, the phrases used in the two test conditions in the Liss et al. (2002) study overlapped. In an attempt to minimize the possibility that listeners simply recognized the lexical targets from the previous condition, rather than benefiting from the target familiarization experience, phrases with low intelligibility were selected for the cross-dysarthria generalization condition. Nonetheless, the limitation of phrase duplication and intelligibility discrepancy remains.

mturk.com).[2] All participants were considered voluntary workers, MTurk workers, protected through MTurk's participation agreement and privacy notice. We utilized the MTurk setup option regarding worker requirements, limiting participation to workers with approval ratings greater than or equal to 99%. We also utilized a location restriction, permitting only workers confirmed as current residents in the United States to be a participant in this study. We also requested that all participants be aged between 19 and 45 years old. If a participant's age fell outside the specified range and/or inclusion criteria were not met, their data were excluded from the study. The use of human subjects recruited via online crowdsourcing was approved by the Utah State University Institutional Review Board.

Speech Stimuli

Test Talker

Speech stimuli used in the current investigation were selected from an extensive database of audio recordings of talkers with dysarthria, collected as part of a larger project conducted in the Motor Speech Disorders Lab at Arizona State University (see Liss et al., 2009 for a full description of speech recording procedures). For the test stimuli, a set of 80 semantically anomalous phrases elicited from a male native talker of American English with ataxic dysarthria secondary to cerebellar disease was used. The talker's speech was characterized perceptually by excess and equal stress, reduced articulation rate, harsh vocal quality, monotone, monoloudness, and imprecise articulation with irregular articulatory breakdowns—all cardinal features of a classic ataxic dysarthria (Darley et al., 1975; Duffy, 2013). The talker's perceptual features have been acoustically validated in a number of earlier studies (Lansford & Liss, 2014a, 2014b; Liss et al., 2009).

The 80 semantically anomalous but syntactically plausible phrases (e.g., *had eaten junk and train*) have been described in greater detail elsewhere (see Liss et al., 2009), but for the purposes of this study, the phrases were carefully designed to reduce the influence of contextual knowledge on intelligibility performance. All phrases were six syllables, alternated in metrical stress, and ranged in length from three to five words. A small subset of 20 phrases served as a brief intelligibility pretest.[3] The remaining 60 phrases served as the posttest.

Familiarization Talkers

For the familiarization stimuli, we selected recordings of passage readings, a 35-phrase adapted version of the standard "Grandfather Passage" (Darley et al., 1975), elicited from seven male native talkers of American English with dysarthria arising from various etiologies and one male native talker of American English with no history of neurological disease or injury (control). Talkers with dysarthria were selected specifically according to similarity of perceptual features to the test talker[4] (same, similar, dissimilar) and level of intelligibility (low, mid, high). Thus, eight familiarization conditions, seven involving a talker with dysarthria and one involving a talker with neurologically healthy speech, were created. A description of the talkers, relative to the similarity and intelligibility designations, is provided in Table 1. To support the similarity designations, which were derived from the perceptual similarity work (see Lansford et al., 2014), Table 1 also highlights a small subset of dysarthria-general and dysarthria-specific perceptual features associated with each talker.[5] The intelligibility designations (low, mid, and high) are based on transcription accuracy data previously reported for these talkers. In particular, those talkers designated as having low, mid, and high intelligibility levels had transcription accuracy scores with ranges of 26%–36%, 47%–56%, and 74%–75%, respectively (see Lansford & Liss, 2014b).

Procedure

A brief description of the study task, requirements (use of headphones), time commitment (approximately 45 min), and remuneration amount ($5) was posted on MTurk. Interested workers were directed to a web page, loaded with a listener perception application hosted on a secure university-based web server. Before beginning the study, MTurk workers were required to read through the consent form approved by the Institutional Review Board and indicate their agreement with the document. Participants were then required to complete a questionnaire regarding age, gender, and basic background questions related to history of impairment and previous experience

[2]We have previously demonstrated that results of perceptual learning with dysarthric speech obtained via MTurk are comparable to those obtained in the controlled laboratory environment (Lansford, Borrie, & Bystricky, 2016).

[3]The pretest was used to yield baseline intelligibility data to determine if listener groups, formed by random assignment to one of eight familiarization conditions, presented with comparable intelligibly levels prior to familiarization. It was, however, important to keep the pretest as brief as possible so that adaptation could largely be attributed to the familiarization condition rather than experience with the test talker during the pretest.

[4]It is tempting to utilize the talkers' etiology-based dysarthria diagnosis as a means for indexing talker similarity for the purposes of this investigation. While this medical model of dysarthria classification offers a convenient and intuitive framework for delineating the varied dysarthria subtypes, not all talkers with the same etiology exhibit similar perceptual features, and perceptual features within a given subtype vary with the severity of speech degradation (see Weismer & Kim, 2010 for a detailed discussion). Recent work done in the area of perceptual similarity of dysarthric speech offers an alternative and ecologically valid approach for investigations that examine sensitivity to shared structure over talkers with dysarthria (Lansford et al., 2014; Lansford, Berisha, & Utianski, 2016).

[5]While beyond the immediate scope of the present analysis, it is important to note that there is abundant evidence of acoustic data to support the ratings of perceptual features provided in Table 1 (see Lansford & Liss, 2014a; Lansford, Liss, & Norton, 2014; Liss et al., 2009; Liss, LeGendre, & Lotto, 2010).

Table 1. Talkers indexed by perceptual similarity and intelligibility level designations, dysarthria diagnosis, and a small subset of dysarthria-general and dysarthria-specific speech features.

				Dysarthria-general		Dysarthria-specific	
Talker	**Similarity**	**Intelligibility**	**Diagnosis**	**Vowel space area**	**Articulation rate**	**Stress rhythm**	**Vocal quality**
1	same	mid	ataxic	reduced	reduced	equal/even	harsh
2	similar	mid	mixed	reduced	reduced	equal/even	harsh
3	similar	high	mixed	reduced	reduced	equal/even	harsh
4	similar	low	ataxic	reduced	reduced	equal/even	harsh
5	dissimilar	mid	ataxic	reduced	reduced	reduced	normal
6	dissimilar	high	mixed	reduced	reduced	equal/even	strained-strangled
7	dissimilar	low	hypokinetic	reduced	reduced	reduced	breathy
8	control	—	—	—	—	—	—

Note. Perceptual similarity designations derived from Lansford, Liss, and Norton (2014). Intelligibility level designations derived from Lansford and Liss (2014b), with low, mid, and high reflecting transcription accuracy scores of 26–36%, 47–56%, and 74–75%, respectively. Talker 8, the control talker, presented with neurologically-healthy speech.

with persons having motor speech disorders. Upon completion of the questionnaire, participants were randomly assigned to one of eight familiarization conditions ($n = 20$), and the experimental portion of the procedure was loaded.

All participants engaged in the same pretest. Participants were informed that they would be presented with short phrases produced by someone with a speech disorder and that while the phrases all contained real English words, they would not necessarily make sense. Phrases were presented one at a time, and following each presentation, participants were instructed to use the keyboard to type out exactly what they thought was being said. Participants were encouraged to guess if unsure. Once they had finished typing their response, participants were prompted to press the return key to move on to the next phrase. The presentation order of the pretest phrases was randomized across participants.

Following the pretest, participants received a familiarization experience using lexical feedback and standard protocols from earlier studies (e.g., Borrie et al., 2012; Liss et al., 2002). Participants were told that they would hear a short passage reading produced by a person with a speech disorder and would see written subtitles of what the talker was saying. They were told that their task was to listen closely and to use the written information to understand what was being said. Participants all heard the same passage reading, but, depending on condition assignment, similarity and intelligibility of the dysarthria differed. After this, all participants engaged in the same posttest, identical in structure to the pretest but using novel testing stimuli (i.e., the posttest phrases).

File Analysis

The total data set consists of 320 transcripts: 160 transcripts of the pretest and 160 transcripts of the posttest. Transcripts were analyzed for a standard measure of speech intelligibility, that is, percent words correct (PWC). Using previously established scoring procedures (Borrie et al., 2012; Liss, Spitzer, Caviness, Adler, & Edwards, 1998), words were counted correct if they matched the intended target or differed only by tense (*-ed*) or plurality (*-s*) and did not add another syllable. Homophones and obvious spelling errors were also counted as correct. Two PWC scores were tabulated for each participant, one for the pretest and one for the posttest. These scores reflect a measure of intelligibility for each participant before and after the familiarization experience. Twenty percent of the transcripts were randomly selected according to computer-generated random number lists and reanalyzed to examine interjudge reliability for coding words correct. Discrepancies revealed high agreement between two independent judges with a Pearson correlation r score above .988.

Data Analysis

Multiple linear regression was used to test whether adaptation to a test talker with ataxic dysarthria, as measured by listener posttest PWC scores, could be predicted by the nature of the listener's prior familiarization experience, specifically similarity of perceptual features (same, similar, dissimilar) and level of intelligibility (low, mid, high). Regression is particularly fitting because it has relatively few assumptions, is useful in testing the effects of grouping variables, and allows the testing of specific levels of each variable. Further, it allows for controlling of the covariates in the model—most importantly the pretest PWC scores. The general regression model was assessed as follows:

$$\widehat{PWC}_{post} = b_0 + b_1\,PWC_{pre} + b_2\,condition, \qquad (1)$$

where $\widehat{PWC}_{post}$ is the predicted PWC scores at posttest, the b's refer to the change in the posttest PWC scores for a one-unit change in the variable, and *condition* refers to the specific experienced combination of similarity and intelligibility. Seven total conditions were compared. Given that the condition variable is categorical, we used "dummy" coding, using the control as the reference condition. The important assumptions regarding normality and heteroscedasticity were evaluated to ensure that a linear modeling scheme was adequate. To end, using linear contrasts, we assessed whether there were differences among the levels of similarity and intelligibility beyond that of the control condition.

Results

Multiple linear regression was used to assess whether similarity and/or intelligibility of the familiarization experience predicted posttest PWC scores, reflecting adaptation to a test talker with ataxic dysarthria. Assumptions of normality and heteroscedasticity were evaluated, confirming that the linear model was appropriate. There were no significant differences at pretest across either similarity or intelligibility conditions ($p = .593$ and $p = .847$, respectively). The full regression model accounted for 70% of the variance in the posttest PWC scores, as shown by the high R^2 value presented in Table 2. It is important to note that even without the pretest PWC scores included in the model, the specific familiarization experience still accounted for 66% of the variance in the posttest PWC scores.

Table 2 presents the regression coefficients in percentage points (i.e., 19.662 for the Same-Mid condition informs that, on average, listeners familiarized with the same talker were 19.662 percentage points higher than listeners familiarized with the control talker). As shown, all combinations of similarity and intelligibility of the familiarization experience significantly predicted higher posttest PWC scores than the control condition. The following results are all controlling for the pretest PWC scores (i.e., each listener is statistically equal at pretest, thus giving us optimally comparable results) even though there were no differences across conditions at pretest.

In Figure 1, we present the means and standard errors of the improvement on PWC from pretest to posttest for each combination of the similarity and intelligibility levels. These results correspond to those reported in the multiple regression model (see Table 2). Of note, the figure suggests there are differences across levels of both similarity and intelligibility. To test these differences (i.e., across levels of similarity and across levels of intelligibility), we used linear contrasts (linear hypotheses). The results of these comparisons are discussed below.

Table 2. Regression results of percent words correct (PWC) posttest scores by the familiarization condition.

Variable	PWC (N = 160)	*SE*	*p*
Pretest	0.319	(0.071)	< .001
Condition			
Control	[reference]	[reference]	[reference]
Same			
Mid	19.662	(1.414)	< .001
Similar			
High	14.424	(1.408)	< .001
Mid	14.172	(1.407)	< .001
Low	11.854	(1.407)	< .001
Dissimilar			
High	4.716	(1.411)	.001
Mid	5.124	(1.408)	< .001
Low	2.975	(1.414)	.037
R^2 of full model		.698	
R^2 without pretest		.657	

Control

Compared with the control condition, listeners familiarized with talkers with similar (13.48 percentage points, $p < .001$) and dissimilar (4.72 percentage points, $p < .001$) perceptual features achieved significantly higher posttest PWC scores, holding the intelligibility level constant. Again, compared with the control condition, listeners familiarized with talkers with low (7.41 percentage points, $p < .001$), mid (not including the same-talker condition; 9.65 percentage points, $p < .001$), and high (9.57 percentage points, $p < .001$) levels of intelligibility significantly predicted higher PWC scores, holding the similarity level constant.

Similarity

We also compared posttest PWC scores across levels of similarity (holding intelligibility constant). Compared with the same-talker condition, listeners familiarized with talkers with similar (−6.18 percentage points, $p < .001$) and dissimilar (−15.39 percentage points, $p < .001$) perceptual features achieved significantly lower posttest PWC scores, holding the intelligibility level constant. The performance following familiarization with talkers with similar and dissimilar perceptual features was also different, where listeners familiarized with similar talkers achieved significantly higher posttest PWC scores (9.21 percentage points, $p < .001$), again holding the intelligibility level constant.

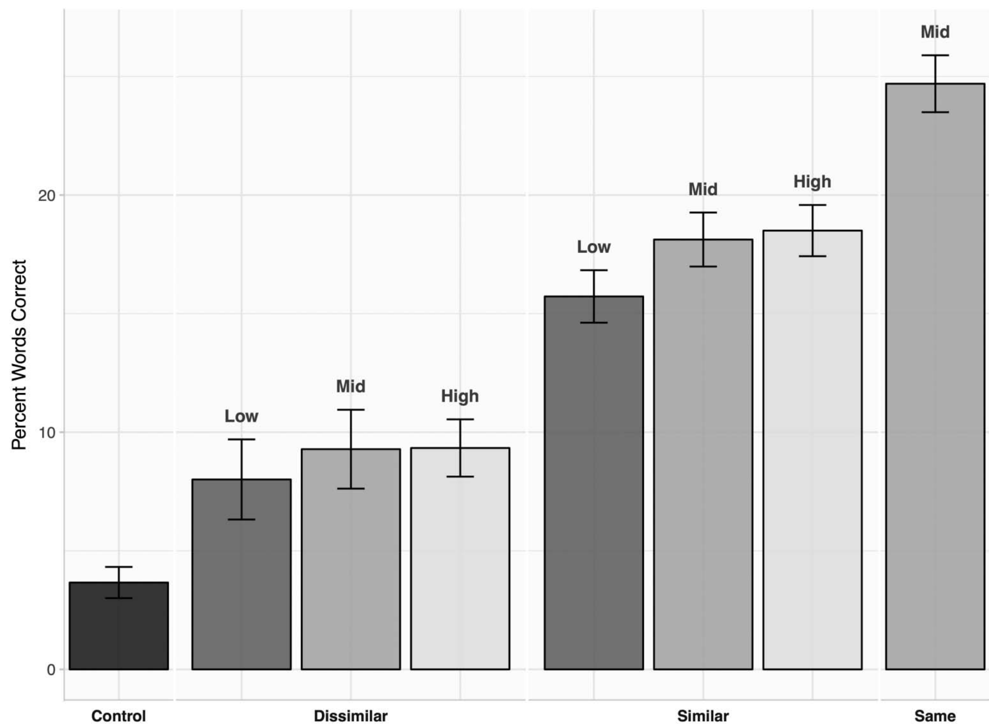
Intelligibility

To end, we compared performance across levels of intelligibility (holding similarity constant). There were differences between the levels of intelligibility, where listeners familiarized with talkers with speech of low intelligibility levels achieved lower posttest PWC scores than listeners familiarized with talkers with mid (−2.23 percentage points, $p = .026$) and high (−2.16 percentage points, $p = .032$) intelligibility levels. There was no significant difference between the mid and high levels of intelligibility ($p = .938$).

Discussion

In this article, we show that perceptual adaptation to a talker with dysarthria can be predicted by the nature of the listener's prior familiarization experience. While all listeners benefited from familiarization with a talker with dysarthria, adaptation was superior when the perceptual features of the familiarization talker were similar to those of the test talker and reduced when the familiarization talker had low intelligibility. Together, these key findings

Figure 1. Perceptual adaptation to a talker with dysarthria by the familiarization condition. Comparison of the improvement on percent words correct from pretest to posttest (y-axis) by similarity of perceptual features (*x*-axis) and level of intelligibility (shade of the bar and labeled). Error bars represent ±1 *SE*.



highlight the differential value of previous experience in adaptation to a novel talker with dysarthria, consistent with central predictions of the ideal adaptor framework (Kleinschmidt & Jaeger, 2015).

The ideal adaptor framework emphasizes statistical learning, whereby listeners utilize talker productions to generate beliefs (and once established, prior beliefs) regarding underlying cue distributions. These prior beliefs can then be used to support perceptual adaptation to a novel talker, or put another way, generalization of perceptual learning across talkers with dysarthria. The framework predicts that prior beliefs are particularly valuable when the distributional regularities in acoustic signal closely match those encountered in a novel talker (i.e., talkers with similar speech patterns). In such cases, listeners require less evidence to establish correct beliefs about the novel talker's speech patterns. On the other hand, when the speech patterns of the novel talker differ substantially from what the listener would expect on the basis of their previous experiences, more evidence is required for adaptation to transpire (Kleinschmidt & Jaeger, 2016). This is precisely what we observed in the current study.

Regardless of similarity of perceptual speech features, listeners familiarized with a talker with dysarthria demonstrated more adaptation than those who were not. This would suggest that listeners are sensitive to distributional regularities in shared structure (i.e., dysarthria-general features) that unify the dysarthrias as a group of motor speech disorders (e.g., reduced vowel space and reduced articulation rate; see Table 1 and Weismer & Kim, 2010). However, we also observed that adaptation to a test talker with dysarthria was superior for listeners familiarized with a talker exhibiting similar perceptual features compared with listeners familiarized with a talker exhibiting dissimilar perceptual features, implicating sensitivity to distributional regularities in specific structure (i.e., dysarthria-specific features). Thus, in support of the predictions of the ideal adaptor framework, the value of prior experience for generalization of learning across talkers with dysarthria is strongly tied to how closely the speech patterns of the familiarization talker match those of the test talker—the more closely matched, the more beneficial the prior beliefs will be (see also Kleinschmidt & Jaeger, 2016, for additional evidence).

When similarity of perceptual features is held constant, adaptation to a talker with dysarthria was significantly reduced for listeners familiarized with dysarthria of low intelligibility relative to listeners familiarized with dysarthria of mid or high intelligibility. No difference for listeners familiarized with dysarthria of mid or high intelligibility was observed. This would suggest some sort of intelligibility threshold for robust perceptual adaptation, whereby acoustic cues are so severely degraded that they restrict the listeners' ability to establish distributional beliefs about the talker's speech patterns. Guediche et al. (2016) have proposed that highly distorted speech signals may violate the process of mapping the incoming auditory input onto stored linguistic representations. Thus, we postulate that the severely degraded acoustic cues that characterize talkers with dysarthria of low intelligibility increases the computational load of distributional learning, constraining adaptation and generalization of learning to a novel talker with dysarthria. Consistent with this hypothesis, other researchers have also documented evidence that more intelligible speech supports greater perceptual adaptation (e.g., Bradlow & Bent, 2008; Guediche et al., 2016).

Clinical Implications

Although largely theoretical, this investigation affords important clinical considerations. While yet to be evaluated in the clinical domain, listener-targeted perceptual learning paradigms, or *perceptual training*, that focus on the primary communication partners of a patient to improve speech intelligibility may be a valuable adjunct to traditional patient-based interventions or a stand-alone alternative when a patient is unable to improve speech behaviorally. Here, we observed that listeners familiarized with the same talker achieved an average of 25% performance gain (see Figure 1). Thus, the current findings add to the body of evidence that support the development and clinical testing of perceptual training paradigms that target the individual's primary communication partner to reduce the intelligibility burden of dysarthria (Borrie et al., 2012; Borrie, McAuliffe, Liss, O'Beirne, & Anderson, 2012; Borrie et al., 2013; Borrie & Schäfer, 2015, 2017; Lansford et al., 2016; Liss et al., 2002).

Another potential clinical application of perceptual learning is use with health care professionals, wherein training may be conducted using dysarthria from specific populations (e.g., amyotrophic lateral sclerosis or Parkinson's disease) or a more general representation of dysarthric speech, depending on the needs of the professional. Here, we observed that listeners familiarized with talkers with similar- and dissimilar-sounding dysarthrias achieved an average of 18% and 9% performance gain, respectively. While certainly not as large as gains achieved following the talker-specific familiarization experience, these generalized adaptations, particularly when prior experience involves a talker with a similar-sounding dysarthria, are clinically meaningful. Thus, the current findings also afford support for the more global applications of perceptual learning paradigms, whereby health care professionals could be trained to better decipher dysarthric speech.

Future Directions

It remains for future research to explicitly identify the content of distributional learning, that is, exactly which general and specific structured regularities in the degraded signal are learned during the familiarization experience. Another important line of work would be to explore if distributional knowledge about a talker with dysarthria is vulnerable to decay over time and, along those lines, whether training-related factors such as variability in familiarization stimuli (i.e., multiple talkers) and amount of familiarization provide an adaptation advantage, particularly when generalization of learning across talkers with dissimilar features is desired. There is certainly evidence that variability in the familiarization experience can benefit adaptation to foreign-accented speech (e.g., Baese-Berk et al., 2013; Bradlow & Bent, 2008). Baese-Berk et al. (2013), for example, observed adaptation to a novel talker with a novel accent (Slovakian-accented English) for listeners familiarized with multiple talkers with different accents (excluding Slovakian-accented English). This result was not evident for listeners familiarized with multiple talkers with the same accent (Mandarin-accented English). The authors comment that exposure to multiple foreign accents of English may better highlight commonalities across foreign accents (e.g., accent-general features such as slow speaking rate). Thus, familiarizing the listener with multiple talkers and a variety of dysarthria presentations may lead to robust adaptation to talkers with dysarthria, more generally.

Conclusion

Perceptual adaptation to a talker with dysarthria can be predicted by the nature of the listener's prior familiarization experience, specifically similarity of perceptual features (same, similar, dissimilar) and level of intelligibility (low, mid, high). While perceptual similarity elevates adaptation and low intelligibility constrains it, listeners still benefit from prior experience with a talker with dysarthria. Taken together, the findings suggest that listeners are sensitive to structure over talkers with dysarthria and can integrate this knowledge, to varying degrees, in subsequent processing of dysarthric speech. Thus, these findings broaden our theoretical knowledge of adaptation to the degraded speech signal, as well as the clinical application of training paradigms that exploit these perceptual processes for therapeutic gain.

Acknowledgments

This research was supported by the National Institute of Deafness and Other Communication Disorders, National Institutes of Health Grant R21 DC 016084, awarded to Stephanie A. Borrie. We gratefully acknowledge Paul Vicioso Osoria and Danielle Fawcett, research assistants in the Human Interaction Lab at Utah State University, for the development of the web-based application and the coding of listener transcripts, respectively. We also extend our gratitude to Julie Liss at Arizona State University for the continued use of her extensive speech sample database.

References

Baese-Berk, M. M., Bradlow, A. R., & Wright, B. A. (2013). Accent-independent adaptation to foreign accented speech. *The Journal of the Acoustical Society of America, 133*(3), EL174–EL180. https://doi.org/10.1121/1.4789864

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,* 290–305. https://doi.org/10.1044/1092-4388

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

Borrie, S. A., McAuliffe, M. J., Liss, J. M., O'Beirne, G. A., & Anderson, T. (2012). A follow-up investigation into the mechanisms that underlie improved recognition of dysarthric speech. *The Journal of the Acoustical Society of America, 132,* EL102–EL108. https://doi.org/10.1121/1.4736952

Borrie, S. A., McAuliffe, M. J., Liss, J. M., O'Beirne, G. A., & Anderson, T. (2013). The role of linguistic and indexical information in improved recognition of dysarthric speech. *The Journal of the Acoustical Society of America, 133,* 474–482. https://doi.org/10.1121/1.4770239

Borrie, S. A., & Schäfer, M. C. M. (2015). The role of somatosensory information in speech perception: Imitation improves recognition of disordered speech. *Journal of Speech, Language, and Hearing Research, 58,* 1708–1716. https://doi.org/10.1044/2015_JSLHR-S-15-0163

Borrie, S. A., & Schäfer, M. C. M. (2017). Effects of lexical and somatosensory feedback on long-term improvements in intelligibility of dysarthric speech. *Journal of Speech, Language, and Hearing Research, 60,* 2151–2158. https://doi.org/10.1044/2017_JSLHR-S-16-0411

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

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign accented English. *The Journal of the Acoustical Society of America, 116*(6), 3647–3658. https://doi.org/10.1121/1.1815131

Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Philadelphia, PA: Saunders.

Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K., & McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: Evidence from the comprehension of noise-vocoded sentences. *Journal of Experimental Psychology: General, 134,* 222–241. https://doi.org/10.1037/0096-3445.134.2.222

Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management* (3rd ed.). St. Louis, MO: Elsevier Mosby.

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics, 36,* 268–294.

Golomb, J. D., Peelle, J. E., & Wingfield, A. (2007). Effects of stimulus variability and adult aging on adaptation to time-compressed speech. *The Journal of the Acoustical Society of America, 121,* 1701–1708. https://doi.org/10.1121/1.2436635

Guediche, S., Fiez, J. A., & Holt, L. L. (2016). Adaptive plasticity in speech perception: Effects of external information and internal predictions. *Journal of Experimental Psychology: Human Perception and Performance, 42*(7), 1048–1059.

Hustad, K. C. (2008). The relationship between listener comprehension and intelligibility scores for speakers with dysarthria. *Journal of Speech, Language, and Hearing Research, 51,* 562–573. https://doi.org/10.1044/1092-4388(2008/040)

Kleinschmidt, D. F., & Jaeger, T. F. (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

Kleinschmidt, D. F., & Jaeger, T. F. (2016). What do you expect from an unfamiliar talker? In J. Trueswell, A. Papafragou, D. Grodner, & D. Mirman (Eds.), *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Lansford, K., Berisha, V., & Utianski, R. (2016). Modeling listener perception of speaker similarity in dysarthria. *The Journal of the Acoustical Society of America, 139*(6), EL209–EL215. https://doi.org/10.1121/1.4954384

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, 25*(2), 233–239. https://doi.org/10.1044/2015_AJSLP-15-0059

Lansford, K. L., & Liss, J. M. (2014a). Vowel acoustics in dysarthria: Mapping to perception. *Journal of Speech, Language, and Hearing Research, 57,* 68–80. https://doi.org/10.1044/10924388(2013/12-0263)

Lansford, K. L., & Liss, J. M. (2014b). Vowel acoustics in dysarthria: Speech disorder diagnosis and classification. *Journal of Speech, Language, and Hearing Research, 57,* 57–67. https://doi.org/10.1044/1092-4388(2013/12-0262)

Lansford, K. L., Liss, J. M., & Norton, R. E. (2014). Free classification of perceptually-similar speakers with dysarthria. *Journal of Speech, Language, and Hearing Research, 57,* 2051–2064. https://doi.org/10.1044/2014_JSLHR-S-13-0177

Liss, J. M., LeGendre, S., & Lotto, A. J. (2010). Discriminating dysarthria type from envelope modulation spectra. *Journal of Speech, Language, and Hearing Research, 53*(5), 1246–1255. https://doi.org/10.1044/1092-4388(2010/09-0121)

Liss, J. M., Spitzer, S. M., Caviness, J. N., & Adler, C. (2002). The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *The Journal of the Acoustical Society of America, 112,* 3022–3030.

Liss, J. M., Spitzer, S., Caviness, J. N., Adler, C., & Edwards, B. (1998). Syllabic strength and lexical boundary decisions in the perception of hypokinetic dysarthric speech. *The Journal of the Acoustical Society of America, 104*(4), 2457–2466. https://doi.org/10.1121/1.423753

Liss, J. M., White, L., Mattys, S. L., Lansford, K. L., Lotto, A. J., Spitzer, S. M., & Caviness, J. N. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research, 52,* 134–1352.

Pallier, C., Sebastián-Gallés, N., Dupoux, E., Christophe, A., & Mehler, J. (1998). Perceptual adjustment to time-compressed speech: A cross-linguistic study. *Memory & Cognition, 26,* 844–851.

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. 229–241). Cambridge, England: Oxford University Press.

Yorkston, K. M., Beukelman, D. R., Strand, E. A., & Bell, K. R. (1999). *Management of motor speech disorders in children and adults*. Austin, TX: Pro-Ed.