Research Article

Regularity Matters: Unpredictable Speech Degradation Inhibits Adaptation to Dysarthric Speech

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

Purpose: Listener-targeted perceptual training paradigms, which leverage the mechanism of perceptual learning, show strong promise for improving intelligibility in dysarthria, offsetting the communicative burden from the speaker onto the listener. Theoretical models of perceptual learning underscore the importance of acoustic regularity (i.e., signal predictability) for listener adaptation to degraded speech. The purpose of the current investigation was to evaluate intelligibility outcomes following perceptual training with hyperkinetic dysarthria, a subtype characterized by reduced signal predictability.
Method: Forty listeners completed the standard 3-phase perceptual training protocol (pretest, training, and posttest) with 1 of 2 talkers with hyperkinetic dysarthria. Perceptual data were compared to a historical data set for 1 other talker

with hyperkinetic dysarthria to examine the effect of perceptual training on intelligibility.
Results: When controlling for pretest intelligibility, regression results suggest listeners of the 2 novel talkers with hyperkinetic dysarthria performed comparably to the listeners of the original talker on the posttest following training. Furthermore, differences between pretest and posttest intelligibility failed to reach clinical significance for all 3 talkers and statistical significance for 2 of the 3.
Conclusion: The current findings are consistent with theoretical models of perceptual learning and suggest that listener adaptation to degraded speech may be negligible for talkers with dysarthria whose speech is marked by reduced signal predictability.

A central goal of dysarthria management is to improve speech intelligibility, defined here as the extent to which a listener understands a speaker's message. The vast majority of interventions targeting reduced intelligibility require the speaker to behaviorally modify their speech to improve the listener's perception (e.g., loud, clear, and slow speech modifications). However, due to the physical and cognitive demands associated with speaker-oriented approaches, not all individuals with dysarthria are appropriate candidates (Duffy, 2013). In response to this critical gap in clinical practice, an alternative approach to intervention, one that targets reduced intelligibility by focusing on the listener rather than the speaker, has been advanced (Liss, 2007). Listener-targeted perceptual training

paradigms in which listeners are familiarized with dysarthric speech show promise for improving intelligibility and enhancing communication in dysarthria, without requiring speaker change (e.g., Borrie, McAuliffe, & Liss, 2012; Lansford, Luhrsen, Ingvalson, & Borrie, 2018; Liss, Spitzer, Caviness, & Adler, 2002).

Perceptual training paradigms leverage the mechanism of perceptual learning to improve listeners' understanding of dysarthric speech. Theoretical models of perceptual learning posit that the familiarization experience affords the listener an opportunity to map degraded, or otherwise noncanonical, acoustic cues, onto linguistic categories stored in memory, resulting in improved perception of that speech in subsequent encounters (Samuel & Kraljic, 2009). Key to successful adaptation to noncanonical speech is the presence of distributional regularities in the speech signal, which arise from both segmental and suprasegmental acoustic cues (Kleinschmidt & Jaeger, 2015). Listeners' knowledge about the distribution of acoustic cues associated with a linguistic category is supported by the statistical predictability of the regularities, thereby driving the cue-to-category mapping process and resulting in improved perception. This learning phenomenon has been well studied with artificially degraded signals including synthetic (e.g.,

[a]School of Communication Sciences and Disorders, Florida State University, Tallahassee
[b]Department of Communicative Disorders and Deaf Education, Utah State University, Logan
[c]Department of Psychology, Utah State University, Logan
Correspondence to Kaitlin L. Lansford: klansford@fsu.edu

Greenspan, Nusbaum, & Pisoni, 1988), noise-vocoded (e.g., Davis & Johnsrude, 2007; Loebach, Bent, & Pisoni, 2008), and time-compressed (Dupoux & Green, 1997) speech and naturally occurring signals, including accented (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Sidaras, Alexander, & Nygaard, 2009), hearing impaired (e.g., McGarr, 1983), and, importantly, dysarthric speech.

Currently, there exists a solid body of evidence suggesting that listener-targeted perceptual training paradigms may be a viable clinical approach for reducing the intelligibility burden associated with dysarthria. Statistically and clinically significant gains to intelligibility, ranging from 8 to 20 percentage point increase in intelligibility from pretest to posttest, have been demonstrated for hypokinetic, ataxic, and spastic dysarthria (Borrie, Lansford, & Barrett, 2017a, 2017b, 2018; Borrie & Schäfer, 2015; Lansford et al., 2018). Notably, these dysarthria subtypes are characterized by rhythmic and phonemic degradations that are largely consistent/stable (e.g., slow rate, equal and even stress, reduced stress, monotone, monoloudness, harsh or breathy vocal quality, imprecise articulation, and reduced vowels). It is presumed that listeners' knowledge of the distribution of acoustic cues associated with a linguistic category is supported by the consistency of these speech degradations. Dysarthric speech, however, is not exclusively characterized by consistent speech features. In cases of neurological disease that result in uncontrolled movement patterns, such as Huntington's disease, the outward flow of speech may be irregularly interrupted or impacted, resulting in the largely inconsistent, and thus unpredictable, speech features characteristic of hyperkinetic dysarthria (e.g., variable speaking rate, excess loudness variations, pitch breaks, inappropriate silences). Reduced signal predictability may also be present in other forms of progressive neurological disease that result in motor instability, which may worsen as the disease progresses (e.g., cerebellar degeneration, Parkinson's disease). If consistent patterns of speech degradations drive perceptual learning in dysarthria, this begs the question: Will the intelligibility benefits associated with perceptual training be diminished, or even nonexistent, for talkers whose speech is characterized by reduced signal predictability?

Although not the original intent, results of a recent study provide preliminary evidence illuminating the role of signal predictability on perceptual learning of dysarthric speech (Borrie et al., 2018). The primary purpose of this work was to examine the role of listeners' rhythm perception abilities relative to intelligibility improvement for two talkers with dysarthria, one whose speech was characterized by largely predictable, but degraded, speech rhythm (hypokinetic dysarthria) and the other whose speech was characterized by unpredictable rhythmic disturbance (hyperkinetic dysarthria). Based on previous findings demonstrating a predictive relationship between rhythm perception and intelligibility improvement following perceptual training with a talker with ataxic dysarthria (Borrie et al., 2017a), it was hypothesized that advanced rhythm perception abilities would support perceptual learning of the talker with

predictably degraded speech rhythm, but not the talker with unpredictable speech rhythm. The results largely supported the hypothesis but also revealed an unexpected finding that none of the listeners, irrespective of rhythm perception abilities, demonstrated improved understanding of the hyperkinetic speaker following familiarization (see Figure 1 for a visual overview of key conclusions from published studies, illustrating significant intelligibility improvements from pretest to posttest for ataxic, spastic, and hypokinetic dysarthria but not for hyperkinetic dysarthria). We speculated that there were simply insufficient distributional regularities present in the speech signal to support the cue-to-category mapping process for this talker during the brief exposure period. However, without replication of this finding in other talkers with inconsistent speech degradations, it remains unknown if the lack of learning was due to reduced signal predictability or to factors that have not yet been considered.

The purpose of the current study was to test this hypothesis, that inconsistent speech degradations inhibit perceptual learning, with two novel talkers with hyperkinetic dysarthria. Using a three-phase perceptual training protocol, 40 naïve listeners were familiarized with one of the two hyperkinetic talkers and their pretest and posttest transcription accuracy scores were measured. The data collected for this project were compared to the historical data, previously reported in Borrie et al. (2018), to determine if listeners undergoing perceptual training perform similarly across three different talkers with inconsistent speech degradations. Demonstration of similar performance across the three unpredictable talkers would indicate that the incidental findings reported in our earlier study are replicable. Next, pretest and posttest transcription accuracy scores for each talker were compared to determine if the familiarization experience leads to statistically and clinically

Figure 1. Key conclusions from published studies, illustrating significant intelligibility improvements from pretest to posttest for ataxic (Borrie et al., 2017a, 2017b; Lansford et al., 2018), spastic (Borrie & Schäfer, 2015), and hypokinetic dysarthria (Borrie et al., 2018), but not for hyperkinetic dysarthria (Borrie et al., 2018).

significant gains to intelligibility. Given the unpredictable nature of the speech features represented in the current study, it was hypothesized that listeners would perform similarly across the three talkers and would not benefit from the familiarization experience. If the hypotheses were supported, the results would lay the foundation for future work to systematically investigate the impact of signal predictability on perceptual learning outcomes in dysarthria.

Method

Listener Participants

Forty adults (20 men, 20 women), aged 18–62 years ($M = 36.4$, $SD = 9.7$), participated in the current study. Participants reported American English as their native language and no history of hearing, speech, language, or cognitive impairment. Furthermore, all participants denied prior significant experience conversing with individuals diagnosed with motor speech disorders. Listener participants were recruited via the crowdsourcing platform Amazon's Mechanical Turk (MTurk; http://www.mturk.com). Briefly, MTurk offers an online labor force in which workers complete small jobs referred to as Human Intelligence Tasks (HITs), in exchange for monetary remuneration. All workers are considered voluntary and are protected through MTurk's participation agreement and privacy notice. Consistent with our previous studies (e.g., Borrie et al., 2018; Lansford, Borrie, & Bystricky, 2016), we required MTurk workers to meet the following qualifications in order to participate: (a) location confirmed in the United States, (b) HIT approval rating of 99% or better, and (c) approval of a minimum of 500 HITs.[1] The assumption is that workers with a 99% approval rating from a minimum of 500 HITs have historically adhered to task instructions. Recruited workers were compensated $5 in exchange for their participation. The institutional review board at Florida State University approved the use of human subjects recruited via MTurk for the current study.

Speech Stimuli

The speech stimuli used for the current investigation were selected from an extensive database of speakers with dysarthria, collected in the Motor Speech Disorders Lab at Arizona State University as part of a larger study (see Liss et al., 2009, for a description of recording procedures). Speech stimuli included audio-recorded productions of a reading passage and a set of 80 semantically anomalous phrases and were produced by two male talkers diagnosed with moderate-to-severe hyperkinetic dysarthria secondary to Huntington's disease, referred to as *HDM3* and *HDM10* throughout this article. The talkers all exhibited the cardinal features of hyperkinetic dysarthria, including variable speaking rate, excess loudness variations, pitch breaks, inappropriate silences, prolonged intervals and phonemes, and irregular articulatory breakdowns.

The speech stimuli were used to create a three-phase perceptual training protocol (pretest, training, and posttest) for each talker. The set of 80 syntactically plausible but semantically anomalous phrases were divided into two smaller subsets and used as stimuli for the pretest (20 phrases) and posttest (60 phrases) transcription tasks (see the Appendix for the full set of phrases). These six-syllable phrases alternated in metrical stress and ranged from three to five words in length (e.g., confused but roared again and mode campaign for budget). The audio recordings of the reading passage were paired with an orthographic transcription and used as stimuli for the training phase of the protocol. The passage was an adapted version of the Grandfather Passage and was composed of 35 phrases, ranging in length from three to 12 words.

Procedure

A HIT was posted to MTurk detailing a description of the task, time commitment (30–45 min), and eligibility criteria. Interested participants were instructed to access the perceptual experiment, hosted on a secure, university-based web server, via an embedded link in the HIT.

The HIT was released in small batches (recruitment restricted to nine participants per batch) to avoid additional MTurk fees and to apply an additional qualifier to prevent participants from completing the task more than once. Once 20 participants completed perceptual training with one talker, the link in the HIT was changed to direct new participants to complete perceptual training with the other talker (20 listeners per talker).

After clicking the link embedded in the HIT, and prior to completing the tasks, participants were instructed to review a consent form approved by the institutional review board and to indicate their consent by clicking the "Agree" button on the screen. Following consent, participants completed a brief demographic survey to denote their age, gender, previous experience with motor speech disorders, and if they had a history of speech, language, hearing, and/or cognitive impairment.

Following completion of the demographic survey, each participant completed a three-phase, talker-specific perceptual training task with one of two talkers with hyperkinetic dysarthria. Task instructions for each phase of the protocol were provided prior to task initiation. First, participants completed a pretest transcription task, consisting of 20 phrases produced by a single talker with hyperkinetic dysarthria. Participants were instructed to listen to each phrase carefully and to type what they heard. They were informed that though the talker had a speech disorder that would make him difficult to understand, they should try their best to transcribe the speech, even if that meant guessing. They were permitted to listen to each phrase only once, and the task was untimed. Immediately following the pretest transcription task, participants underwent the training phase in which they were familiarized with the

---

[1]Requesters must approve HITs completed by workers before monetary compensation is disbursed.

same talker heard during the transcription pretest. Participants were instructed to listen to the talker's production of each phrase of the Grandfather Passage, while simultaneously following along with the orthographic transcription presented on the screen. The passage phrases were presented one at a time, and participants were instructed to advance to the next phrase when ready. Finally, listeners completed a posttest transcription task in which they were asked to listen to and transcribe 60 novel phrases produced by the same talker heard in the prior two phases. The same task instructions provided at pretest were reiterated at posttest.

Transcript Analysis

Pretest and posttest listener transcripts were scored for words correct using Autoscore, an open-source, computer-based tool for automated scoring of transcripts (http://autoscore.usu.edu; Borrie, Barrett, & Yoho, 2019).[2] Autoscore has scoring rules that can be selected, depending on the needs of the project. Here, we used rules to score words as correct if they match the intended target exactly or differed only by tense or plurality. Homophones and obvious spelling errors were scored as correct using a pre-programed "default" list of common misspellings. A percent words correct (PWC) score was tabulated for the pretest and posttest experimental phases, resulting in a pretest PWC score and a posttest PWC score for each listener, by talker condition.

Data Analysis

To determine if listeners undergoing perceptual training would perform similarly across multiple talkers with unpredictably degraded speech, the perceptual data collected for the two talkers with hyperkinetic dysarthria (HDM3 and HDM10) were compared to a historical data set collected from 50 listeners via MTurk for a third talker with hyperkinetic dysarthria (hereafter referred to as HDM8), described in full detail in Borrie et al. (2018).[3] Importantly, all three talkers exhibited the cardinal features of hyperkinetic dysarthria, including variable speaking rate, excess loudness variations, pitch breaks, inappropriate silences, prolonged intervals and phonemes, and irregular articulatory breakdowns. Although it is possible that intelligibility outcomes could be differentially affected by the unique speech features present in the learning material arising from different talkers, we contend that the talkers used in this study are sufficiently similar to permit comparison of outcomes across the talker set.

To assess whether there were differences in intelligibility improvement across the three talkers, we used linear regression with the posttest PWC score predicted by talker controlling for pretest PWC score. In essence, this compares the intelligibility improvement scores following perceptual training (i.e., comparing posttest after making all individuals statistically equal at pretest). With the three talker groups, we used the previously collected talker data, HDM8, as the reference category and used a linear contrast to compare the two new talkers, HDM3 and HDM10. It is worth noting that inclusion of HDM8's data permits a well-powered comparison. Since a null result is anticipated here (i.e., listeners will perform similarly across the three talkers), inclusion of the historical data allows us to derive as precise a null result as possible, given the current study design. Lastly, we assessed whether there was any improvement in intelligibility for either HDM3 or HMD10. To do so, we used paired t tests and reported the standardized effect sizes of the improvement.

Results

The regression results (reported in Table 1) suggest neither HDM3 nor HDM10 was significantly different than HDM8 ($p = .111$ and $p = .860$, respectively) in terms of posttest PWC when controlling for pretest PWC. The differences between the groups were all small. The two novel talkers, HDM3 and HDM10, were also compared using a linear contrast. The difference was not significant ($p = .325$). All differences between talkers in the sample were small with the standardized coefficients (adjusted standardized mean differences) between −.027 and .114 in comparison to the reference HDM8 talker.

The regression results are supported by visual analysis of the distribution of intelligibility improvement scores (i.e., difference between posttest and pretest intelligibility, reported as percentages) across the three talkers (presented in Figure 2). Intelligibility improvement, or lack thereof, for each group of listeners is illustrated by the area below each curve. As shown, the change observed within each talker ranges from approximately 10 percentage point

---

[2]Autoscore has been validated as an accurate (99% accuracy) and efficient scoring tool on both in-house and independent data sets (Borrie et al., 2019).

[3]The methods used to collect the historical data set are described in full detail in Borrie et al. (2018); however, it is important to note that the data collection methods, including experimental phrase lists, used in the present analysis are identical to those used in the earlier study. It is, therefore, appropriate to draw comparisons between these data sets.

Table 1. Linear regression results showing the unstandardized coefficients, the 95% confidence intervals (CIs), the standardized coefficients, and their associated *p* values.

Variable	Estimate	95% CI Lower	95% CI Upper	Standardized estimate	*p* value
Talker					
HDM8	[ref]	[ref]	[ref]	[ref]	[ref]
HDM3	1.786	−0.417	3.991	0.140	.111
HDM10	−0.350	−4.289	3.588	−0.027	.860
Pretest	0.877	0.754	1.000	0.920	< .001
Intercept	6.655	0.137	13.174	−0.025	.046

Note. HDM3 and HDM10 were not significantly different based on a linear contrast (*p* = .325). [ref] = linear regression reference category.

Figure 2. The distributions of intelligibility improvement for each listener within each talker. The area under each curve represents the density of responses at each value of improvement (e.g., most listeners for HDM10 had improvement between 2 and 4).
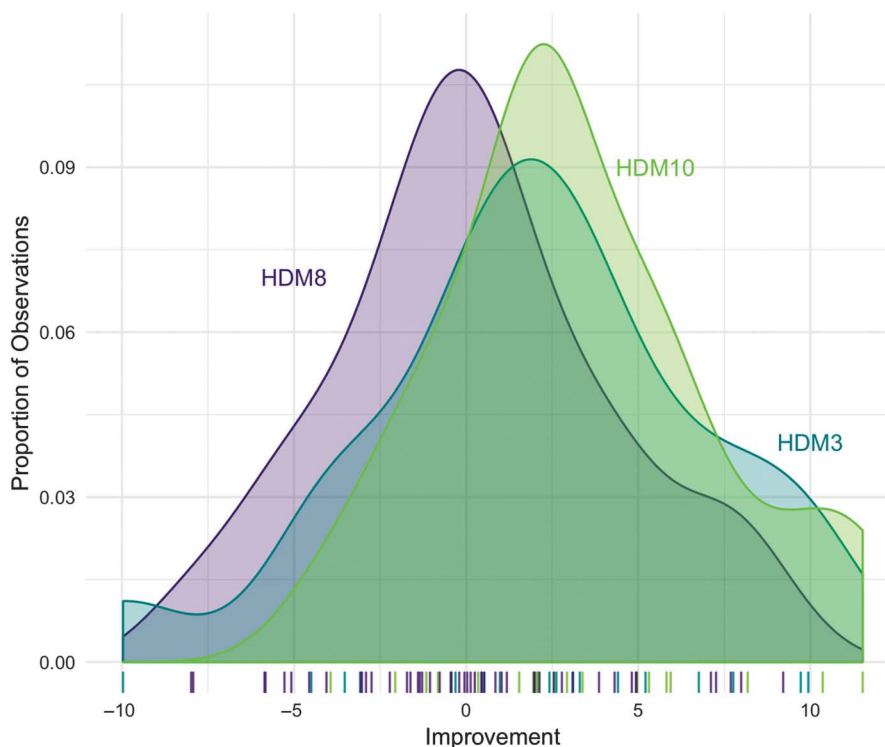
decrease to 10 percentage point improvement. The overall distributions of improvement are similar across talkers, with most listeners only improving between 0 and 5 percentage points.

Figure 3 presents the average pretest and posttest PWC scores for each talker across the listeners with their associated standard errors. The results of the paired t test analyses and the standardized effect sizes are shown in Table 2 and suggest that only the listeners of HDM10 significantly improved from pretest to posttest, although the improvement was very small (3.15 percentage points, on average; $p = .002$). Notably, HDM10 had lower average intelligibility than both HDM3 and HMD8 at both pretest and posttest.

Discussion

In a previous report, we found that listeners derived no perceptual benefit following familiarization with a single talker diagnosed with hyperkinetic dysarthria secondary to Huntington's disease (Borrie et al., 2018). This unexpected result deviated from previous findings that have consistently demonstrated both clinically and statistically significant intelligibility improvements following familiarization with dysarthric speech. We speculated that listeners failed to adapt to the hyperkinetic talker's speech due to reduced signal predictability, arising from involuntary movements

that inconsistently interrupt speech production. The current investigation sought to test this hypothesis by evaluating perceptual outcomes following familiarization with two additional hyperkinetic talkers whose speech was also characterized by inconsistent speech degradations (e.g., variable rate, pitch breaks, excess loudness variation, inappropriate

Figure 3. The average intelligibility scores, indexed by percent words correct, for both pretest and posttest for each talker. The error bars represent $SE = ±1$.
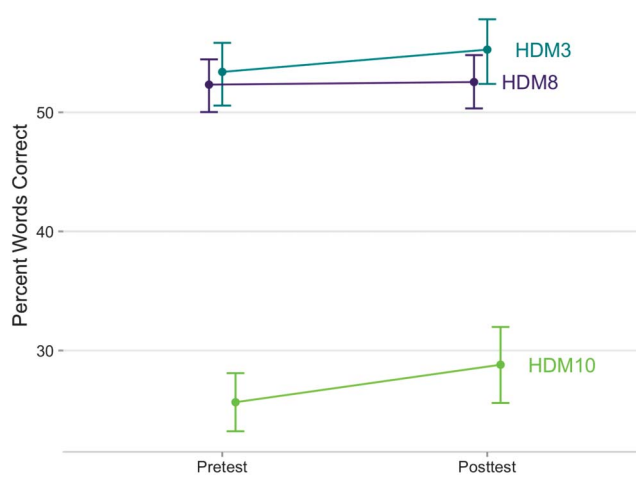
Table 2. Results of the paired-samples *t* tests and the standardized effect sizes.

Talker	Posttest–pretest	*t* statistic	*p* value	Standardized ES
HDM8	0.22	0.374	.710	0.027
HDM3	1.87	1.71	.104	0.297
HDM10	3.15	3.55	.002	0.476

pauses). The current results largely support our original speculation regarding the value of signal predictability in perceptual learning of dysarthric speech. When controlling for pretest intelligibility scores, the results of the regression analysis suggest that listeners of the two novel talkers (HDM3 and HDM10) performed comparably to the listeners of the original talker (HDM8) on the posttest intelligibility task following familiarization, thereby replicating our prior incidental findings. Furthermore, comparisons of pretest to posttest PWC failed to demonstrate statistically significant differences for two of the talkers, who notably had equivalent pretest intelligibility (HDM8 and HDM3).

The difference between pretest and posttest PWC (approximately 3 percentage point increase at posttest) was statistically significant for HDM10, the most severe of the three talkers. The clinical and theoretical implications of this finding, however, should be interpreted with caution. First, the experimental design did not include a true control condition, in which listeners are familiarized with healthy speech, but tested on dysarthric speech. This is not an insignificant caveat. In a recent investigation, listeners assigned to a control condition in which they were familiarized with healthy speech but tested on dysarthric speech gained about 5 percentage points to PWC at posttest for a talker with ataxic dysarthria, relative to significantly higher PWC gains in seven different conditions involving familiarization with dysarthric speech (Borrie et al., 2017b). These findings suggest that some learning likely transpires during the posttest task itself, but that intelligibility improvement is optimized following familiarization with dysarthric speech. Thus, in this study, we cannot be certain that the statistically significant pretest to posttest increase demonstrated for listeners of HDM10 would be significantly different from a control condition involving healthy speech. Furthermore, although the 3–percentage point difference is statistically significant, its effect size is small and it fails to reach the threshold of a clinically significant change to intelligibility, considered by some to be between 5% and 8% for sentence level intelligibility (Stipancic, Tjaden, & Wilding, 2016; Yorkston, Beukelman, & Traynor, 1984). Additionally, the results of the regression analysis indicated that the intelligibility improvement revealed for HDM10 was not significantly different than the pretest to posttest differences revealed for HDM8 and HDM3, which, recall, both failed to reach statistical significance. Lastly, when compared to previously reported gains of 8–20 percentage points posttraining for listeners of talkers diagnosed with other, more predictably degraded dysarthria subtypes, the current

findings suggest intelligibility gains were, at a minimum, reduced for talkers whose speech is marked by unpredictable acoustic cues. Thus, not only are these findings consistent with theoretical models of perceptual learning, they also suggest that perceptual training may not be a viable treatment option for talkers with dysarthria whose speech is marked by inconsistent speech features. Additional questions, posed below, must first be addressed before this clinical implication can be firmly concluded.

The traditional three-phase perceptual training paradigm utilizes a rather brief familiarization task in which listeners hear 35 phrases totaling a single-passage reading, produced by a talker with dysarthria, and use externally provided lexical feedback (i.e., orthographic transcriptions of the audio phrases) to map the degraded acoustic cues onto linguistic categories stored in the memory. It is plausible that this task is not optimal for facilitating cue-to-category mapping when the signal is marked by substantial unpredictability. It may be the case, then, that listeners would benefit from simply more exposure to unpredictable speech. Alternatively, there is some evidence that listeners might benefit varied exposure to dysarthric speech, as the perceptual benefits associated with perceptual training do not appear to be solely talker-specific. Rather, listeners demonstrate generalized adaptation to dysarthric speech in which training with one talker improves intelligibility of a novel talker (Borrie et al., 2017b). Though the magnitude of intelligibility improvement associated with generalized adaptation appears to be related to the degree of the perceptual feature overlap between the training and test talkers, training with any talker with dysarthria led to greater intelligibility improvements than training with a healthy control speaker in our earlier study. In other words, listeners benefit not only from the presence of distributional regularities specific to a talker but also to those common across talkers with dysarthria. Additional support for this conclusion comes from a recent study that demonstrated listeners trained and tested on similarly accented speech outperformed those trained and tested on accented speech that was perceptually dissimilar on a transcription task (Alexander & Nygaard, 2019). Thus, perhaps exposure to varied talkers with more predictably degraded speech features would permit generalizable cue-to-category mapping, resulting in improved intelligibility of less predictable speech. This is an empirical question that warrants attention.

The use of externally provided lexical feedback is considered to be an integral component of the traditional three-phase perceptual training paradigm; however, intelligibility outcomes might be optimized with the addition of internally generated somatosensory feedback, via a vocal imitation task, during the familiarization experience. Using a speaker with spastic dysarthria as a test case, listeners provided with both lexical and somatosensory feedback achieved significantly greater gains to intelligibility relative to listeners who were provided with only lexical feedback during the familiarization experience (Borrie & Schäfer, 2015). It was postulated that somatosensory feedback aids

in disambiguating the dysarthric signal, thereby facilitating cue to category. Future work should evaluate whether intelligibility outcomes following perceptual training with dysarthria characterized by largely unpredictable speech degradation could be improved with the provision of additional levels of feedback during training.

Although hyperkinetic dysarthria presents a convenient test case for examining the effects of signal unpredictability on learning outcomes, it is unlikely that these findings are exclusive to hyperkinetic speech. Rather, we argue that intelligibility outcomes will be diminished for any degraded speech signal characterized by inconsistent speech features, as their presence drives down the predictability of the acoustic cues. The current results lay the groundwork for future investigations to systematically evaluate the impact of specific speaker parameters, including signal predictability, on the magnitude of intelligibility improvement following perceptual training. However, in order to comprehensively examine the effects of signal predictability on learning outcomes in dysarthria, it will become necessary not only to develop and validate methods for quantifying acoustic and perceptual predictability but also to evaluate which unpredictable speech features (temporal and/or spectral) are most deleterious to learning.

Finally, as is the case for any study that examines treatment-related changes using group averages, this work is limited with regard to examining individual variability in learning outcomes. In the current study, the average intelligibility change across the three speakers from pretest to posttest was approximately 1.25 percentage points. Notably, though, some listeners improved intelligibility by more than 10 percentage points, whereas others deteriorated at posttest (see Figure 2). Given that perceptual training targets the listener, it will be critically important for future work to evaluate how listener-related parameters (e.g., rhythm perception, age, hearing acuity, and other cognitive factors) interact with speaker-related parameters (e.g., signal predictability and overall severity of the speech disorder) during the familiarization process.

Conclusions

Listener-targeted perceptual training paradigms that improve intelligibility of dysarthric speech show strong potential, offsetting the communicative burden from the speaker onto the listener. As perceptual training moves closer to clinical implementation, the consideration of candidacy for this potential treatment option is imperative. Our previous work has demonstrated robust intelligibility gains following a simple perceptual training paradigm for a variety of talkers with dysarthria whose speech is largely characterized by predictable, albeit disordered, features. Results of this study, however, suggest that intelligibility gains following perceptual training may be negligible for talkers with dysarthria whose speech is largely characterized by unpredictable features. The current results call for rigorous study of the role of signal predictability in perceptual learning of dysarthric speech, which is of significant clinical importance and informs theoretical models regarding mechanisms of learning.

References

Alexander, J. E. D., & Nygaard, L. C. (2019). Specificity and generalization in perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America, 145*(6), 3382–3398. https://doi.org/10.1121/1.5110302

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). 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. (2017b). Generalized adaptation to dysarthric speech. *Journal of Speech, Language, and Hearing Research, 60*, 3110–3117. https://doi.org/10.1044/2017_JSLHR-S-17-0127

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*, 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., & Schäfer, M. C. (2015). The role of somatosensory information in speech perception: Imitation improves recognition of disordered speech. *Journal of Speech, Language, and Hearing Research, 58*(6), 1708–1716. https://doi.org/10.1044/2015_JSLHR-S-15-0163

Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to nonnative speech. *Cognition, 106*, 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*, 3647–3658. https://doi.org/10.1121/1.1815131

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: Top-down influences on the interface between audition and speech perception. *Hearing Research, 229*(1), 132–147. https://doi.org/10.1016/j.heares.2007.01.014

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

Dupoux, E., & Green, K. (1997). Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 914–927. https://doi.org/10.1037/0096-1523.23.3.914

Greenspan, S. L., Nusbaum, H. C., & Pisoni, D. B. (1988). Perceptual learning of synthetic speech produced by rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 421–422.

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

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., Luhrsen, S., Ingvalson, E., & 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,* 91–98. https://doi.org/10.1044/2017_AJSLP-17-0090

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. 18–219). San Diego, CA: Plural.

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. https://doi.org/10.1121/1.1515793

Liss, J. M., White, L., Mattys, S. L., Lansford, K. L., Spitzer, S., Lotto, A. J., & Caviness, J. N. (2009). Quantifying speech rhythm abnormalities in the dysarthrias. *Journal of Speech, Language, and Hearing Research, 52*(5), 1334–1352. https://doi.org/10.1044/1092-4388(2009/08-0208)

Loebach, J. L., Bent, T., & Pisoni, D. B. (2008). Multiple routes to the perceptual learning of speech. *The Journal of the Acoustical Society of America, 124,* 552–561. https://doi.org/10.1121/1.2931948

McGarr, N. S. (1983). The intelligibility of deaf speech to experienced and inexperienced listeners. *Journal of Speech and Hearing Disorders, 26*(3), 451–458. https://doi.org/10.1044/jshr.2603.451

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics, 71*(6), 1207–1218. https://doi.org/10.3758/APP.71.6.1207

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,* 3306–3316. https://doi.org/10.1121/1.3101452

Stipancic, K. L., Tjaden, K., & Wilding, G. (2016). Comparison of intelligibility measures for adults with Parkinson's disease, adults with multiple sclerosis, and healthy controls. *Journal of Speech, Language, and Hearing Research, 59,* 230–238. https://doi.org/10.1044/2015_JSLHR-S-15-0271

Yorkston, K. M., Beukelman, D. R., & Traynor, C. (1984). *Assessment of intelligibility of dysarthric speech.* Austin, TX: Pro-Ed.

Appendix

Set of Semantically Anomalous Phrases Used for the Pretest and Posttest Transcription Tasks

account for who could knock
address her meeting time
admit the gear beyond
advance but sat appeal
afraid beneath demand
amend estate approach
and spoke behind her sin
appear to wait then turn
assume to catch control
attack became concerned
attend the trend success
avoid or beat command
award his drain away
balance clamp and bottle
beside a sunken bat
bolder ground from justice
bush is chosen after
butcher in the middle
career despite research
cheap control in paper
commit such used advice
confused but roared again
connect the beer device
constant willing walker
cool the jar in private
darker painted baskets
define respect instead
distant leaking basement
divide across retreat
done with finest handle
embark or take her sheet
for coke a great defeat
forget the joke below
frame her seed to answer
functions aim his acid
had eaten junk and train
her owners arm the phone
hold a page of fortune
increase a grade sedate
indeed a tax ascent

its harmful note abounds
kick a tad above them
listen final station
mark a single ladder
mate denotes a judgment
may the same pursued it
measure fame with legal
mistake delight for heat
mode campaign for budget
model sad and local
narrow seated member
or spent sincere aside
pain can follow agents
perceive sustained supplies
pick a chain for action
pooling pill or cattle
push her equal culture
rampant boasting captain
remove and name for stake
resting older earring
rocking modern poster
rode the lamp for teasing
round and bad for carpet
rowing farther matters
seat for locking runners
secure but lease apart
signal breakfast pilot
sinking rather tundra
sparkle enter broken
stable wrist and load it
submit his cash report
support with dock and cheer
target keeping season
technique but sent result
thinking for the hearing
to sort but fear inside
transcend almost betrayed
unless escape can learn
unseen machines agree
vital seats with wonder