JSLHR

Research Article

When Additional Training Isn't Enough: Further Evidence That Unpredictable Speech Inhibits Adaptation

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

Purpose: Robust improvements in intelligibility following familiarization, a listener-targeted perceptual training paradigm, have been revealed for talkers diagnosed with spastic, ataxic, and hypokinetic dysarthria but not for talkers with hyperkinetic dysarthria. While the theoretical explanation for the lack of intelligibility improvement following training with hyperkinetic talkers is that there is insufficient distributional regularity in the speech signals to support perceptual adaptation, it could simply be that the standard training protocol was inadequate to facilitate learning of the unpredictable talker. In a pair of experiments, we addressed this possible alternate explanation by modifying the levels of exposure and feedback provided by the perceptual training protocol to offer listeners a more robust training experience.

Method: In Experiment 1, we examined the exposure modifications, testing whether perceptual adaptation to an unpredictable talker with hyperkinetic dysarthria could be achieved with greater or more diverse exposure to dysarthric speech during the training phase. In Experiment 2, we examined feedback modifications, testing whether perceptual adaptation to the unpredictable talker could be achieved with the addition of internally generated somatosensory feedback, via vocal imitation, during the training phase.

Results: Neither task modification led to improved intelligibility of the unpredictable talker with hyperkinetic dysarthria. Furthermore, listeners who completed the vocal imitation task demonstrated significantly reduced intelligibility at posttest.

Conclusion: Together, the results from Experiments 1 and 2 replicate and extend findings from our previous work, suggesting perceptual adaptation is inhibited for talkers whose speech is largely characterized by unpredictable degradations. Collectively, these results underscore the importance of integrating signal predictability into theoretical models of perceptual learning.

Familiarization paradigms in dysarthria offer a promising platform for listener-targeted remediation by exploiting the malleability of the listeners' perceptual system for the express purpose of improving intelligibility of the disordered speech signal. Briefly, familiarization paradigms involve an exposure experience, in which feedback may or may not be provided, to train listeners with an individual's specific speech pattern. Theories of perceptual learning posit that experience with degraded, or otherwise noncanonical, speech facilitates mapping of the acoustic cues onto linguistic categories stored in memory (i.e., cue-to-category mapping). According to Kleinschmidt and Jaeger's ideal adaptor framework (2015), the cue-to-category mapping process is fostered by the distributional regularities present in the speech signal, arising from both segmental and suprasegmental acoustic cues. Listeners' knowledge of the distribution of these cues is supported, then, by the statistical predictability of the regularities, leading to the formation of generative models of the noncanonical speech that aid in perception of that speech during subsequent encounters. Thus, at least theoretically, noncanonical speech patterns that are rich in predictable acoustic information should be highly learnable, thereby offering a clear path to perceptual adaptation. Indeed, laboratory findings appear to uphold such theoretical models of perceptual learning, in which improved perception of noncanonical speech following familiarization has been revealed for both artificially generated (i.e., time-compressed and noise-vocoded speech; Davis & Johnsrude, 2007; Dupoux & Green, 1997; Loebach et al., 2008) and naturally occurring speech signals (e.g.,

[a]School of Communication Science & 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

Editor-in-Chief: Bharath Chandrasekaran
Editor: Kate Bunton
Received December 5, 2019
Revision received February 21, 2020
Accepted February 24, 2020
https://doi.org/10.1044/2020_JSLHR-19-00380

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

foreign accented and dysarthric speech; Borrie et al., 2017a, 2017b; Bradlow & Bent, 2008; Liss et al., 2002; Sidaras et al., 2009). Understanding the role of signal predictability in perceptual learning of noncanonical speech was not a central focus of the aforementioned studies, but it bears noting that the artificially generated and naturally occurring speech signals utilized in these previous studies were largely composed of predictable acoustic information.

There has been a recent shift in focus, however, to understanding the role of signal predictability on intelligibility of dysarthric speech following familiarization (Borrie et al., 2018; Lansford et al., 2019). Presently, there exists a substantial body of evidence demonstrating improved understanding of talkers diagnosed with hypokinetic, ataxic, and spastic dysarthria following familiarization, ranging from 8- to 20-percentage-point increases at posttest (Borrie et al., 2017a, 2017b, 2018; Borrie & Schäfer, 2015, 2017; Kim, 2015; Lansford et al., 2016; Lansford et al., 2018; Liss et al., 2002).[1] In general, these dysarthria subtypes are characterized by largely consistent segmental and suprasegmental degradations (e.g., slow rate, equal and even stress, reduced stress, monotone, monoloudness, harsh or breathy vocal quality, imprecise articulation, and reduced vowels), offering listeners an adequate level of signal predictability to support perceptual adaptation. Dysarthria, however, is not exclusively characterized by predictable acoustic degradation. In fact, neurological impairments that cause reduced motor stability and/or involuntary movement of the speech mechanism often result in inconsistent, and therefore unpredictable, speech degradations, including variable speaking rate, pitch, and loudness; inappropriate silences; and irregular articulatory breakdown (Duffy, 2013; Spencer & Dawson, 2019; see Table 1 for a list of common segmental and suprasegmental features of dysarthria categorized by level of predictability). Given the presumed importance of signal predictability to perceptual learning theory (Kleinschmidt & Jaeger, 2015), it follows that intelligibility improvement following familiarization with dysarthric speech should be constrained by the constellation of predictable and unpredictable speech features present in the speech signal.

While intelligibility outcomes associated with familiarization with dysarthric speech have yet to be systematically investigated relative to the level of signal predictability afforded by the signal, there is preliminary evidence suggesting perceptual learning does not transpire for talkers whose dysarthria is largely characterized by unpredictable speech features (Borrie et al., 2018; Lansford et al., 2019). Though not the original intent of the investigation, we recently found that listeners familiarized with a talker with hypokinetic dysarthria, characterized by predictably disordered speech rhythm, demonstrated improved perception of that speech at posttest, but listeners of a talker with hyperkinetic dysarthria, with unpredictably disordered speech rhythm, derived no perceptual benefit from familiarization (Borrie et al., 2018). We hypothesized 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. This hypothesis was later supported by an investigation designed to explicitly examine perceptual learning outcomes for two additional talkers with unpredictable speech secondary to hyperkinetic dysarthria, in which clinically significant gains to intelligibility following perceptual training were not observed (Lansford et al., 2019). It is important to consider that individuals diagnosed with hyperkinetic dysarthria are particularly vulnerable to reduced signal predictability due to the underlying involuntary movement of the musculature responsible for speech production. Thus, replication of negligible learning following familiarization with two additional talkers with unpredictable speech secondary to hyperkinetic dysarthria presents compelling evidence underscoring the importance of signal predictability to perceptual learning paradigms in dysarthria. This point is additionally strengthened when we contrast these null results with the robust improvements in intelligibility demonstrated for talkers with more predictably degraded speech (e.g., Borrie et al., 2017a, 2017b; Borrie & Schäfer, 2015, 2017; Kim, 2015; Lansford et al., 2016; Lansford et al., 2018).

Though the theoretical explanation for the lack of clinically and statistically significant changes to intelligibility following training with the hyperkinetic talkers was that there was insufficient distributional regularity in their speech signals to support cue-to-category mapping during the training phase, it may simply be the case that, for talkers whose speech is largely characterized by unpredictable degradations, the standard training protocol is insufficient to facilitate learning. Indeed, it is widely accepted, across disciplines, that behavioral outcomes are modifiable based on the task used to elicit them. It is because of this universally recognized phenomenon that we have come to understand some of the conditions that optimize familiarization of dysarthric speech. In the first study to systematically examine the nature of the familiarization experience on intelligibility outcomes in dysarthria, Borrie, McAuliffe, Liss, Kirk, et al. (2012) found that the provision of lexical feedback during the training phase (in the form of an orthographic transcript paired with the incoming speech signal) was integral to the cue-to-category mapping process. This seminal finding, replicated in a follow-up study (Borrie, McAuliffe, Liss, O'Beirne & Anderson, 2012), not only duplicated results revealed for perceptual learning of noise-vocoded (Davis & Johnsrude, 2007) and foreign-accented speech (Bradlow & Bent, 2008) but also guided the standardization process of the perceptual training protocol used in our subsequent studies (see Borrie et al., 2017a, 2017b, 2018; Lansford et al., 2019; Lansford et al., 2016; Lansford et al., 2018). Briefly, the standard protocol for talker-specific training is composed of three phases: pretest, training, and posttest. During the pretest and posttest phases, listeners complete orthographic transcription tasks, which provide

[1]Operational definitions of "clinically significant" gains to intelligibility in dysarthria ranging from 5% to 12% (Stipancic et al., 2016, 2018; Van Nuffelen et al., 2010).

Table 1. Common segmental and suprasegmental speech errors in dysarthria, categorized by assumed impact on overall signal predictability.

Speech feature	Segmental errors	Suprasegmental errors
Predictable	Vowel distortions, imprecise consonants, characterized by distortions and undershoot	Consistently too slow or consistently too fast, monopitch, monoloudness, equal and even stress, reduced stress
Unpredictable	Irregular articulatory breakdown	Variable speaking rate, short rushes of speech, pitch breaks, excess loudness variations, irregular pauses

stable estimates of intelligibility for a talker, both before and after training. During the training phase, participants listen to a reading passage produced by the talker and follow along with an orthographic transcript of the passage (i.e., lexical feedback). In general, a change in intelligibility from pretest to posttest, beyond what is revealed for a control condition, is taken as evidence of talker-specific adaptation resulting from the training experience. It bears noting that, while lexical feedback is provided by the standard protocol, the familiarization experience is quite brief. If the largely unpredictable speech signal offers any learnable acoustic information (i.e., distributional regularities), it is plausible that perceptual adaptation could be achieved with a more robust training experience.

Given the unpredictable nature of hyperkinetic dysarthria, one could argue that the current training experience simply does not afford listeners with enough exposure to the degraded speech to facilitate learning of the largely unpredictable signal. Although robust improvements in intelligibility have been revealed for talkers with more predictably degraded speech following the standard familiarization task, perhaps the cue-to-category mapping process for a talker whose speech is largely characterized by unpredictable speech degradations is not well supported by the brevity of this task. Perceptual learning of noncanonical speech is driven by the listener's knowledge of the distributional regularities present in the signal, and acquisition of this knowledge requires experience with the signal. It follows, then, that more experience with a degraded signal should enhance perception of that signal in subsequent encounters. Empirical support for this argument comes from literature documenting perceptual learning of spectrally rotated (Green et al., 2013) and time-compressed (Banai & Lavner, 2016) speech, in which learning outcomes were optimized with more exposure to the degraded speech signal. In addition, listeners familiarized with talkers with spastic and athetoid dysarthria, secondary to cerebral palsy, over multiple sessions, demonstrated gradual improvements in consonant identification (Kim, 2015). Thus, if the unpredictable talker's speech is at all learnable, one option for modifying the training task is to provide the listeners with simply more exposure during the training experience with that talker's speech.

Generalized adaptation paradigms, in which listeners are trained with one talker but tested on a novel talker, offer another opportunity for providing listeners with more exposure during the training task, as intelligibility improvement associated with perceptual training is not solely talker specific. Rather, listeners can generalize perceptual training effects to novel, untrained talkers. Such generalized adaptation to novel, untrained talkers has been revealed for foreign-accented (Alexander & Nygaard, 2019; Xie & Myers, 2017), noise-vocoded (Huyck et al., 2017), and dysarthric (Borrie et al., 2017b) speech. The magnitude of intelligibility improvement associated with generalized adaptation is constrained by degree of perceptual similarity between the training and test talkers, such that training with a perceptually similar talker results in greater intelligibility improvement than training with a perceptually dissimilar talker (Alexander & Nygaard, 2019; Borrie et al., 2017b). These results are consistent with the underlying theoretical assumptions of the ideal adaptor framework: Listeners are sensitive to a shared structure over talkers and similar situations and can draw on previous experience to support subsequent processing. While the dysarthria subtypes are characterized by unique constellations of perceptual features (i.e., dysarthria-specific features), Weismer and Kim (2010) argue that, in general, the dysarthrias are unified by a subset of perceptual features (i.e., dyarthria-general features), including reduced vowel space, reduced speaking rate, reduced formant transitions and articulatory speed, and reduced phonetic contrasts. We have observed that listeners can leverage the acoustic regularity afforded by these largely consistent speech features, common to most talkers with dysarthria, to support not only their learning of a trained talker but also for novel talkers who share this acoustic structure (Borrie et al., 2017b). Thus, if the acoustic structure of the unpredictable talker with hyperkinetic dysarthria overlaps at all with the acoustic structure of dysarthric speech, more generally, listeners may be able to generalize knowledge gained through familiarization with dysarthric talkers with more predictable, and therefore learnable, speech to improve understanding of a dysarthric talker whose speech is largely characterized by unpredictable speech degradations.

Alternatively, one could argue that the current training experience does not afford listeners with enough or, alternatively, the right type of feedback to facilitate learning of disordered speech that is largely characterized by unpredictable speech degradations. Recall that, during training, listeners hear the talker read aloud a single reading passage (35 phrases) while following along with an orthographic transcription. This provision of lexical feedback is considered to be essential to robust perceptual learning of dysarthric speech (Borrie, McAuliffe, Liss, Kirk, et al., 2012; Borrie, McAuliffe, Liss, O'Beirne, & Anderson, 2012). Speech acquisition and production models (e.g., Directions Into Velocities of Articulators; Guenther et al., 1998), however, implicate the importance of somatosensory feedback for guiding perceptual learning of a teaching signal, considered a requisite and interconnected cognitive process of motor learning. According to such models, when a listener encounters a teaching signal, a neural network is activated to aid in recognition (i.e., cue-to-category mapping). Key to this model is that speech production attempts are essential for mapping the acoustic input onto linguistic representations. Thus, theoretically, vocal imitation facilitates both perceptual and motor learning of a teaching signal by linking the acoustic target with the somatosensory information required to produce that sound. Indeed, recent evidence linking the provision of somatosensory feedback to improved visual perception of vowels lends some support to this model assumption (Masapollo & Guenther, 2019). Furthermore, Borrie and Schäfer (2015) found that listeners who received both lexical and somatosensory feedback, via a vocal imitation task, during familiarization with a talker with spastic dysarthria achieved significantly greater intelligibility gains than listeners who were provided with lexical feedback alone. Additionally, listeners who made imitation attempts that more closely matched the dysarthric target achieved greater gains to intelligibility, providing additional support to underlying theoretical assumptions. A follow-up study with ataxic dysarthria revealed that the use of somatosensory feedback during training was required to facilitate long-term retention of intelligibility gains (Borrie & Schäfer, 2017). Overall, then, learning outcomes were optimized for listeners of talkers with spastic and ataxic dysarthria with the provision of somatosensory feedback. Thus, it is plausible that if the largely unpredictable speech signal offers any learnable acoustic information, the provision of somatosensory feedback could support perceptual adaptation to that speech.

The Current Study

The theoretical explanation for the lack of intelligibility improvement following training with hyperkinetic talkers in our earlier study is that there was insufficient distributional regularity in the speech signals to support perceptual adaptation (Lansford et al., 2019). However, it is possible that, despite the overall unpredictable nature of the acoustic cues present in this talker with hyperkinetic dysarthria, there may be some distributional regularities in the speech that could be exploited for learning if given a more robust training experience. To address this possibility, we conducted two independent experiments, using existing literature to support modifications to the training phase of the standard three-phase perceptual training protocol. In the first experiment (Experiment 1), we examined the modification of *more and different exposure,* testing whether perceptual adaptation to an unpredictable talker could be achieved with either repeated exposure to the test talker with hyperkinetic dysarthria or with exposure to talkers with dysarthria who present with more predictably degraded speech during the training phase. In the second experiment (Experiment 2), we examined *more and different feedback,* testing whether perceptual adaptation to an unpredictable talker could be achieved with the addition of internally generated somatosensory feedback, via vocal imitation, during the training phase. To address these research questions, we examined the listeners' pretest and posttest transcripts to determine if there was a clinically and statistically significant increase in word accuracy following the training experience. Although both exposure and feedback training modifications have resulted in optimized learning outcomes in work examining intelligibility of talkers with dysarthria characterized by more consistent speech degradations, we hypothesize that learning outcomes in this study will be diminished, or even negligible, due to the lack of acoustic predictability afforded by the speech signal.

Method

Experiment 1: Exposure Modification

Listener Participants

Forty adult listeners (14 females), ranging from 24 to 59 years, participated in Experiment 1. Listener participants were recruited via Amazon Mechanical Turk (MTurk; http://www.mturk.com). MTurk, a crowdsourcing platform, offers researchers a convenient and efficient option for collection of behavioral data. In our previous work, we have observed not only data equivalence between laboratory participants and MTurk workers (Lansford et al., 2016) but also statistically and clinically significant gains in intelligibility following perceptual training with dysarthric speakers (Borrie et al., 2017b, 2018). Briefly, workers complete online tasks, 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. To participate in this study, we required that MTurk workers meet the following qualifications: (1) location confirmed in the United States; (2) HIT approval rating of 99% or better (rating provided by MTurk to characterize worker HIT performance); (3) completion of a minimum of 500 HITs; (4) no history of speech, language, or hearing disorders; (5) native speaker of American English; (6) no significant experience communicating with individuals with dysarthria; and (7) use of headphones during the perceptual experiment. Recruited workers were compensated $7 in exchange for their participation. The Institutional Review Board at Florida State University approved the use of human subjects recruited via MTurk for this study.

Talkers

Audio recordings of speech produced by three male talkers with dysarthria were used for Experiment 1 (see Table 2 for talker descriptions and Liss et al., 2009, for details related to the speech recording procedures). The hyperkinetic talker was the same unpredictable talker used in a previous study, in which a group of 50 listeners

Table 2. Talker demographics and perceptual features characteristic of the presenting dysarthria.

Talker demographics	Dysarthria subtype and etiology	Severity rating	Perceptual symptoms present	Dysarthria-general perceptual symptoms present
Male, 43 years	Hyperkinetic dysarthria secondary to Huntington's disease	Moderate-to-severe	• Prolonged intervals and phonemes • Variable rate (but reduced more generally) • Inappropriate silences • Irregular articulatory breakdown • Intermittent hypernasality	• Reduced speaking rate • Reduced vowel space area • Reduced formant transitions • Imprecise articulation
Male, 84 years	Ataxic dysarthria secondary to cerebellar degeneration	Moderate	• Excess and equal stress • Prolonged phonemes and intervals • Monotone • Monoloudness • Irregular articulatory breakdown	
Male, 46 years	Mixed spastic-flaccid dysarthria secondary to amyotrophic lateral sclerosis	Moderate	• Prolonged syllables • Monotone • Monoloudness • Hypernasality	

Note. Presenting perceptual features and severity ratings have been perceptually and acoustically validated in earlier studies (Lansford & Liss, 2014a, 2014b; Liss et al., 2009).

recruited via mTurk, who underwent the traditional perceptual training paradigm with lexical feedback, showed no evidence of intelligibility improvement (Borrie et al., 2018). Perceptual training with the other two talkers with more predictably degraded dysarthria (i.e., ataxic and mixed spastic–flaccid dysarthria) yielded intelligibility improvements of 16–20 percentage points across a series of previous studies utilizing both laboratory and crowd-sourced participants (Borrie et al., 2017a, 2017b; Lansford et al., 2018).

Speech Stimuli

The speech stimuli for the pretest and posttest phases consisted of 80 syntactically plausible but semantically anomalous phrases produced by the male talker with hyperkinetic dysarthria. Each phrase is composed of six syllables with alternating syllabic stress, ranging from three to five words in length (e.g., "admit the gear beyond" and "distant leaking basement"). The stimuli used for the training conditions came from recordings of an adapted version of the "Grandfather Passage," produced by the three talkers with dysarthria described in Table 1. The passage reading contains 35 phrases or sentences, ranging from three to 12 words. These stimuli have been described in greater detail in our previous work (Borrie et al., 2017a, 2017b, 2018; Lansford et al., 2019; Lansford et al., 2016; Lansford et al., 2018).

Procedures

A HIT, detailing a description of the task, estimated time commitment (30–45 min), and eligibility criteria were posted to MTurk. The perceptual experiment, hosted on a secure university-based web server, was embedded in a link on the HIT. After clicking the link embedded in the HIT and prior to completing the perceptual experiment, participants were asked to review a consent form approved by the Institutional Review Board and to indicate their consent by clicking the "Agree" button on the screen. Participants who clicked "Agree" were then asked to complete a brief demographic survey to denote their age, sex, 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, the participants were randomly assigned to one of the extended-exposure training conditions, *lexical + repeated* or *lexical + general* (n = 20 listeners per condition). Participants then completed a three-phase, extended-exposure, perceptual training paradigm (pretest, training, posttest). First, all listeners, irrespective of training condition, were asked to transcribe 20 phrases produced by the test talker with hyperkinetic dysarthria. In the task instructions, listeners were informed that the talker would be difficult to understand but was producing real English words, and they should try their best to transcribe his speech even if they were merely guessing. They were also instructed to listen carefully as they would have only one opportunity to hear each phrase.

Following the pretest transcription task, listeners underwent an extended-exposure familiarization task, in which they heard either the talker with hyperkinetic dysarthria produce the reading passage three times (*lexical + repeated* condition) or three talkers with dysarthria, including two talkers with predictably degraded speech and the test talker with hyperkinetic dysarthria, produce the reading passage (*lexical + general* condition). Externally provided lexical feedback (i.e., orthographic transcript) accompanied the recorded productions. Listeners were instructed to listen to each recording, while following along with the orthographic transcript.

Immediately following the extended-exposure training task, listeners completed the posttest, in which they transcribed 60 novel phrases produced by the same talker with hyperkinetic dysarthria from the pretest phase.[2]

Data Analysis

Pretest and posttest listener transcripts were scored for words correct using Autoscore[3] (http://autoscore.usu.edu; Borrie et al., 2019), an open-source computer-based tool for automated scoring of transcripts. This tool has a number of scoring rules that can be selected; we used rules that tell Autoscore to score words as correct if they match the intended target exactly or differ only by tense or plurality. Homophones and obvious spelling errors were also scored as correct. A percent words correct (PWC) score was generated for the pretest and posttest, resulting in a pretest PWC score and a posttest PWC score for each listener.

To assess changes in PWC from pretest to posttest, we used a simple paired-sample *t* test for each condition. Furthermore, we used linear regression with posttest PWC as the outcome predicted by condition, controlling for pretest PWC, to assess for any differences in learning across conditions. This approach provides comparisons between the two conditions while accounting for differences in listener's initial ability to understand the talker. All analyses were performed in R Version 3.6.1 (R Core Team, 2019) in RStudio (RStudio Team, 2018) using ggplot2, dplyr, data.table, janitor, rio, and broom packages (Chan et al., 2018; Dowle & Srinivasan, 2019; Firke, 2019; Robinson & Hayes, 2019; Wickham, 2016; Wickham et al., 2019).

Experiment 2: Feedback Modification

Listener Participants

Forty-nine adult listeners (all female), ranging from 18 to 25 years of age, participated in Experiment 2. Our participant recruitment strategy differed for Experiment 2, in that we chose to collect the data in the laboratory rather than via MTurk. This methodological decision was made to facilitate monitoring of task instruction adherence during the vocal imitation training task. Participants were recruited from a Communication Science and Disorders Research Evaluation class at Florida State University. To be eligible to participate, we required listeners to be native speakers of American English with no self-reported history of hearing, speech-language, or cognitive impairments. While these participants had some background knowledge of dysarthria from introductory coursework, they were not experienced listeners. Participants were compensated with class credit for their participation. The Institutional Review Board at Florida State University approved the use of human subjects for this study.

Talker and Speech Stimuli

Experiment 2 used audio recordings of speech stimuli produced by the same male talker diagnosed with hyperkinetic dysarthria secondary to Huntington's disease, described in greater detail in Experiment 1. The stimuli used for pretest and posttest transcription tasks consisted of the same 80 syntactically plausible but semantically anomalous phrases as described in Experiment 1. The training task stimuli came from the same passage reading as described in Experiment 1, adapted from the Grandfather Passage.

Procedures

The perceptual data were collected in the Motor Speech Disorders Lab at Florida State University. Each participant attended a single session, in which the tasks took approximately 60 min to complete. Prior to initiating the experimental tasks, listeners were asked to review a written description of the experiment and accompanying consent form. They were given an opportunity to ask any questions. After signing the consent form, listeners completed a brief demographic survey to indicate their age, sex, 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, the participants were randomly assigned to one of two experimental conditions: *somatosensory feedback* (n = 25) or *somatosensory + lexical feedback* (n = 24).

A three-phase, vocal-imitation, perceptual training paradigm (pretest, vocal imitation task, posttest) was used for this study. The pretest and posttest transcription tasks across the two training conditions were identical and mirrored those described in Experiment 1. Listeners were seated in front of a desktop computer and fitted with headphones. Task instructions for the pretest and posttest transcription tasks were provided by the research assistant prior to initiation of the tasks (described in detail in Experiment 1). Briefly, listeners were asked to listen carefully to each phrase produced by a talker with disordered speech, as they would be provided only one opportunity to hear each phrase, and to write what they heard in the text box provided on the screen using the keyboard.

Immediately following the pretest transcription task, all listener participants completed a vocal imitation task. The experimental protocol was delivered through the desktop computer. Listeners were fitted with a high-quality earset microphone (Countryman E6 Omnidirectional microphone) to record their imitations for future analysis (not reported here). The level of lexical feedback provided to the participants was manipulated during the vocal imitation task, resulting in two feedback conditions: (1) *lexical + somatosensory feedback condition,* in which the listeners were provided the orthographic transcription of the phrases being produced by the speaker and were asked to imitate the speaker's production of that phrase, and (2) *somatosensory feedback condition,* in which the participants heard

[2]This traditional familiarization paradigm utilizes 20 phrases in the pretest to ensure a stable baseline measure of intelligibility while minimizing the risk of familiarizing listeners to the test talker at pretest. The posttest utilizes 60 phrases.

[3]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) and thus reliability measures for scoring the transcripts in this study were not deemed necessary.

the phrase, unaccompanied by the orthographic transcription, and were asked to imitate the speaker's production. As part of the task instructions, the research assistant modeled the desired behavior using a sentence produced by the hyperkinetic talker. The participants were able to go through the stimuli at their own pace and were permitted to revise their imitations, but were provided only one opportunity to hear each phrase. The research assistant went into a room adjacent to the laboratory to monitor completion of the vocal imitation task out of sight of the participant. This methodological decision was made to help reduce any potential performance anxiety experienced by the participant. Immediately following the vocal imitation task, listeners completed the posttest transcription task (described above).

Data Analysis

Pretest and posttest listener transcripts were scored for words correct using Autoscore (described in full detail under Experiment 1). A PWC score was generated for the pretest and posttest, resulting in a pretest PWC score and a posttest PWC score for each listener, per feedback condition. The same statistical analyses were performed for Experiment 2 as was done for Experiment 1. That is, we examined the change in PWC from pretest to posttest for each condition using paired-sample *t* tests. In addition, we used linear regression with posttest PWC as the outcome predicted by condition, controlling for pretest PWC, to assess for any differences in learning across conditions. As in Experiment 1, all analyses were performed in R and RStudio using the same packages.

Results

Experiment 1

As illustrated in Figure 1 and according to the *t* tests, the PWC scores from the *lexical + repeated exposure* condition did not significantly change from pretest to posttest (mean change = −2.63, $t[19] = -2.05$, $p = .054$). A similar pattern was seen for the *lexical + general exposure* condition (mean change = −2.71, $t[19] = -1.81$, $p = .086$). These findings are comparable to the traditional training paradigm (*lexical*), with the same talker, in which we previously observed no adaptation following training (Borrie et al., 2018; data included in Figure 1 for reference). According to the regression model, there were no differences between the conditions after controlling for pretest (adjusted mean difference = 0.136, $p = .939$).

Experiment 2

As illustrated in Figure 2 and according to the *t* tests, both the *somatosensory* and *lexical + somatosensory* conditions showed significant decreases in PWC from pretest to posttest (see Figure 2). The *somatosensory only* condition decreased by 8.8 percentage points ($t[24] = -10.7$, $p < .001$), while the *lexical + somatosensory* condition decreased by 7.1 percentage points ($t[23] = -6.1$, $p < .001$). The results of Experiment 2 indicate that the provision of somatosensory feedback during the training task did not lead to improved intelligibility of the talker with hyperkinetic dysarthria. In fact, participants in both somatosensory feedback conditions performed significantly worse at posttest than pretest. Furthermore, according to the regression model, there were no differences between the conditions after controlling for pretest PWC (adjusted mean difference = 0.433, $p = .653$).

Discussion

Intelligibility improvement following familiarization has been revealed for talkers diagnosed with spastic, ataxic, and hypokinetic dysarthria (Borrie et al., 2017a, 2017b, 2018; Borrie & Schäfer, 2015, 2017; Lansford et al., 2016; Lansford et al., 2018), but not hyperkinetic dysarthria

Figure 1. The average intelligibility scores, indexed by percent words correct, for both pretest and posttest for the *lexical + repeated* and *lexical + general* conditions in Experiment 1, with the error bars representing ± 1 *SE*. The *lexical* condition shown is historic data regarding this speaker, from Borrie et al. (2018), and is included as reference.

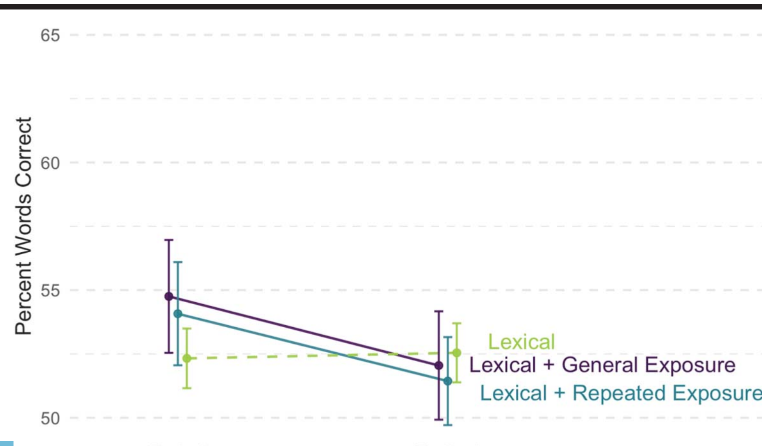
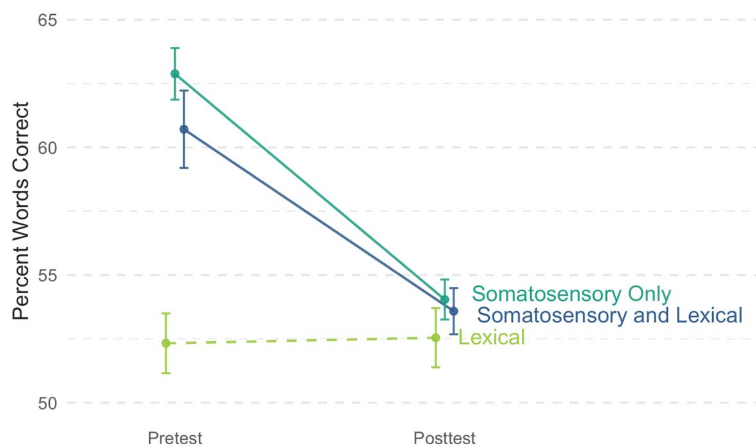


Figure 2. The average intelligibility scores, indexed by percent words correct, for both pretest and posttest for each intelligibility (percent words correct) for the *somatosensory* and *somatosensory + lexical* conditions in Experiment 2, with the error bars representing ± 1 *SE*. The *lexical* condition shown is historic data regarding this speaker, from Borrie et al. (2018), and is included as reference.



(Borrie et al., 2018; Lansford et al., 2019). Although the dysarthria subtypes are collectively characterized by degraded segmental and suprasegmental acoustic information, the underlying neurological impairment constrains the regularity, or predictability, of the speech degradations. The aforementioned empirical findings, then, are consistent with theoretical models of perceptual learning, such that learning outcomes are diminished (or nonexistent) if the training signal is characterized by reduced acoustic regularity. Given that perceptual outcomes are sensitive to task-related parameters, the primary purpose of the current work is to determine if our previous findings are indeed consistent with model predictions, or if they were merely the result of an inadequate training task. To achieve this, we modified the training task to provide either more/different exposure to dysarthric speech (Experiment 1) or more/different feedback regarding the unpredictably disordered speech signal (Experiment 2). These specific modifications were selected because they have been demonstrated to optimize learning outcomes of dysarthric speech (e.g., Borrie & Schäfer, 2015, 2017; Kim, 2015). If listeners benefitted by the provision of more/different information during the training task, it would challenge our postulation that the distributional cues in this hyperkinetic speech sample are entirely unpredictable. Taken together, then, the results of Experiments 1 and 2 suggest the unpredictable training signal, and not the training task, is responsible for negligible learning outcomes for the speaker with hyperkinetic dysarthria.

In Experiment 1, we found that intelligibility of the hyperkinetic talker was not improved with the provision of more exposure (three times more than our traditional training phase) to the unpredictable talker during the perceptual training experience. Thus, even when listeners are provided with additional exposure to the unpredictable talker, listeners did not benefit from the familiarization experience. Likewise, more exposure to dysarthric talkers whose speech is characterized by greater acoustic regularity did not improve intelligibility of the hyperkinetic talker, suggesting that even if listeners were sensitive to dysarthria-general distributional regularities present in the training stimuli, they were unable to apply this information to the unpredictable talker with dysarthria. Thus, it appears that while the test talker's speech signal is characterized by some dysarthria-general perceptual features, there is likely insufficient overlap between the acoustic structure of these general features in the predictable training talkers and the hyperkinetic talker to support any sort of generalized adaptation. Our position that the distributional cues associated with this talker with hyperkinetic dysarthria truly are unpredictable is, therefore, supported by these results.

In Experiment 1, listeners were provided two additional opportunities to hear the same training material. Therefore, while these listeners were afforded with more exposure to the training material during the familiarization period relative to the traditional perceptual training phase (see Borrie et al., 2018), the linguistic material used to facilitate the cue-to-category mapping process was limited to a single reading passage. This raises the possibility that more exposure to unpredictable speech coupled with more linguistically diverse material may provide listeners a greater opportunity to learn. However, for this to be true, the largely unpredictable signal would have to offer at least some level of acoustic regularity to support perceptual adaptation.

It currently remains unknown if listeners of an unpredictable talker would benefit from perceptual training extended over multiple training sessions. Previous studies have, indeed, demonstrated gradual improvement in perception of synthetic (Fenn et al., 2003), spectrally rotated (Green et al., 2013), time-compressed (Banai & Lavner, 2016), and foreign-accented (Earle & Myers, 2015; Xie et al., 2018) speech, over training sessions that spanned multiple days. Listeners from these experiments, then, benefitted not only from more exposure to the noncanonical speech but also from offline, or sleep-mediated, consolidation, a cognitive process hypothesized to support integration of the training experience with prior linguistic knowledge (McClelland et al., 1995). The influence of offline consolidation on intelligibility outcomes has not been explicitly examined in dysarthria; however, gradual improvement over multiple training sessions (Kim, 2015) and long-term retention of perceptual gains posttraining have been previously demonstrated for talkers with more consistently degraded dysarthrias (Borrie, McAuliffe, Liss, Kirk, et al., 2012; Borrie & Schäfer, 2017). Though we strongly suspect unpredictable speech diminishes perceptual learning effects, offline consolidation relative to perceptual training outcomes, more generally, should be considered in future work.

In Experiment 2, we found that intelligibility of the hyperkinetic speaker was not improved with the provision of somatosensory feedback during the perceptual training experience. In fact, listeners across both somatosensory feedback conditions performed worse at posttest (more on this point later). Recall, speech acquisition and production models underscore the importance of production attempts to learning a novel speech sequence. Specifically, by linking the auditory speech sound with its corresponding tactile-kinesthetic (i.e., somatosensory) reference frame, these attempts are hypothesized to support mapping of the resulting acoustic input onto stored linguistic representations. Empirical support for this assumption has been revealed for listeners of a talker with spastic dysarthria, in which the additional provision of somatosensory feedback during the training experience led to greater intelligibility improvement than the provision of lexical feedback alone (Borrie & Schäfer, 2015). However, if we once again return to theoretical models of perceptual learning, intelligibility improvement would not be expected, even with the provision of additional somatosensory feedback, when the novel speech pattern lacks acoustic regularity, as is the case for the hyperkinetic talker. The current findings support this hypothesis.

Interestingly, while the provision of somatosensory feedback via vocal imitation of talkers with spastic and ataxic dysarthria and foreign-accented speech has been demonstrated to improve listeners' intelligibility outcomes relative to the provision of lexical feedback alone (Adank et al., 2010; Borrie & Schäfer, 2015, 2017), work examining perceptual learning of nonnative contrasts found diminished perceptual outcomes when the training task included a production component (Baese-Berk & Samuel; 2016; Leach & Samuel, 2007). Baese-Berk and Samuel (2016) argued that engaging in the vocal imitation task increased the cognitive load experienced by the listener, resulting in reduced learning of the nonnative contrast. This reasoning may actually explain our unexpected findings of significantly reduced intelligibility outcomes at posttest for listeners who imitated the hyperkinetic talker during the training experience. In contrast to earlier vocal imitation studies in which listeners imitated dysarthric talkers with largely predictable patterns of speech degradations (Borrie & Schäfer, 2015, 2017), we assume that imitating the speech of an unpredictable talker was cognitively demanding. Rather than implementing a stable production pattern (e.g., slow speech rate + vocal strain for the talker with spastic dysarthria) to novel phrases, the participants in this study had to adopt a novel pattern for each phrase due to the inconsistent nature of the talker's speech degradations. We suspect that the increased cognitive demands required to complete this task led to fatigue, resulting in the reduced intelligibility at posttest (Hornsby et al., 2016). This conjecture is supported anecdotally by participant feedback. Many participants reported that the vocal imitation task was very challenging and were surprised to learn that the speech stimuli were produced by the same talker. Thus, we speculate that the reduced intelligibility of the hyperkinetic talker following vocal imitation was a transient consequence of listener fatigue induced by the cognitively demanding vocal imitation training task. Given the current findings and the known detrimental effects of increased cognitive load on perceptual outcomes (e.g., Hornsby et al., 2016; Hunter & Pisoni, 2018), it will be important to explicitly examine the extent to which intelligibility outcomes are constrained by the cognitive demands associated with the perceptual training task.

While not a target of this investigation, the findings of Experiment 2 revealed that the listeners who were recruited to complete the vocal imitation task had higher pretest intelligibility scores than the listeners recruited via MTurk in Experiment 1. It is worth reiterating that in order to monitor task adherence during the vocal imitation conditions, we collected the perceptual data in the lab and utilized convenience samples of undergraduate students in the Communication Science and Disorders major. It is very likely that these samples, composed of individuals with introductory knowledge of communication disorders, but with limited experience with disordered speech, were more motivated and, perhaps, demonstrated greater listener effort than the typical sample drawn from MTurk. Given that perception of degraded speech requires increased listening effort to accurately map the acoustic information onto stored linguistic representations (e.g., Peelle, 2018; Pichora-Fuller et al., 2016), and that this effort can be dampened by high levels of acoustic degradation, it is likely that intelligibility of dysarthric speech is influenced by listener effort. If learning had been revealed for the samples of listeners used in Experiment 2, it would be difficult to ascertain its exact cause (training task modification vs. differing listener characteristics). Learning, however, was not observed. Thus, the absence of learning outcomes following training with an unpredictable talker may be robust to differing levels of listener effort and motivation. Nevertheless, given that listener-related parameters, such as rhythm perception ability, vocabulary level, age, and hearing acuity have been shown to mediate speech perception and adaptation (Borrie et al., 2017b, 2018; Ingvalson et al., 2017a, 2017b; McAuliffe et al., 2013), future work should investigate how other listener factors, including motivation and effort and feedback preference (Lametti et al., 2012), impact intelligibility outcomes associated with perceptual training.

Together, the results from Experiments 1 and 2 replicate and extend findings from our previous work, suggesting unpredictable speech inhibits perceptual adaptation during familiarization (Borrie et al., 2018; Lansford et al., 2019). Although the current results provide additional evidence supporting expert perceptual judgments that the speech signal from this talker lacks the acoustic regularity required to support perceptual learning, there are several important factors (many of them already mentioned) that should be considered relative to the results. Importantly, although diminished learning outcomes associated with familiarization has been replicated in three talkers with unpredictably degraded dysarthrias (see Lansford et al., 2019), the current results are limited to a single test talker. These results, therefore, should be interpreted with caution. Despite this limitation, the current results, together with our earlier work (Borrie et al., 2018; Lansford et al., 2019), underscore the importance of integrating signal predictability into theoretical models of perceptual learning. We contend that signal predictability is unlikely a binary construct (predictable vs. unpredictable), but it rather exists on a continuum, tracking to both temporal and spectral acoustic degradations in the speech signal. Although the talkers examined in this study and in our earlier work had hyperkinetic dysarthria, we argue that the level of signal predictability is not constrained by dysarthria subtype. Rather, we expect that any talker with neurological impairments that cause reduced motor stability and/or involuntary movement of the speech mechanism is vulnerable to reduced signal predictability. Furthermore, we expect that the level of signal predictability in talkers with such motoric impairments may vary as a consequence of overall severity of their impairment. Thus, it will be imperative to establish measurement procedures to accurately quantify the level of signal predictability present in a speech signal to support rigorous examination of this novel perceptual construct (i.e., signal predictability) in a diverse cohort of talkers with dysarthria. Only then, would we be able to understand how intelligibility outcomes associated with perceptual training are mediated by the level of predictability afforded by a speech signal.

Clinical Implications

Listener-targeted perceptual training paradigms offer a promising platform for ameliorating intelligibility disorders associated with dysarthria without requiring speaker change. Clinical implementation of perceptual training would address a critical gap in clinical practice that disproportionately affects clinical populations who are not suitable candidates for traditional, speaker-oriented, behavioral treatment (e.g., clear and loud speech modifications) due to significant neuromuscular and/or cognitive impairment (Duffy, 2013). As perceptual training moves closer to clinical implementation, it is important to consider candidacy for this potential treatment approach. Our recent and current findings suggest perceptual training may not be a viable option for individuals whose speech is largely characterized by unpredictable speech degradations. While hyperkinetic dysarthria presented a convenient test case for examining the role of signal predictability in perceptual learning outcomes, we do not propose that candidacy decisions should be made on the basis of dysarthria subtype diagnosis alone. Rather, we argue that this decision should be guided by the level of predictability available in the speech signal.

Acknowledgments

This research was supported by the National Institute on Deafness and Other Communication Disorders through Grant R21DC016084, awarded to Stephanie Borrie. We extend our gratitude to Julie Liss at Arizona State University for the continued use of her extensive speech sample database. Data collection for this project was part of two undergraduate honors theses completed at Florida State University by Cassidy Flechaus and Jenna Wolff. We also acknowledge research assistants, Aja Centrone, Audrey Hendrix, and Madison Morgan, for their assistance with data collection.

References

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science, 21*(12), 1903–1909. https://doi.org/10.1177/0956797610389192

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

Baese-Berk, M. M., & Samuel, A. G. (2016). Listeners beware: Speech production may be bad for learning speech sounds. *Journal of Memory and Language, 89,* 23–36. https://doi.org/10.1016/j.jml.2015.10.008

Banai, K., & Lavner, Y. (2016). The effects of exposure and training on the perception of time-compressed speech in native versus nonnative listeners. *The Journal of the Acoustical Society of America, 140*(3), 1686–1696. https://doi.org/10.1121/1.4962499

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*(11), 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*(5), EL379–EL385. https://doi.org/10.1121/1.5037620

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. J. (2012). A follow-up investigation into the mechanisms that underlie improved recognition of dysarthric speech. *The Journal of the Acoustical Society of America, 132*(2), EL102–EL108. https://doi.org/10.1121/1.4736952

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*(6), 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*(8), 2151–2158. https://doi.org/10.1044/2017_JSLHR-S-16-0411
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
Chan, C.-H., Chan, G. C. H., Leeper, T. J., & Becker, J. (2018). *rio: A Swiss-army knife for data file I/O*. R Package Version 0.5.16.
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
Dowle, M., & Srinivasan, A. (2019). *data.table: Extension of 'data.frame.'* R Package Version 1.12.6. https://CRAN.R-project.org/package=data.table
Duffy, J. R. (2013). *Motor speech disorders: Substrates, differential diagnosis, and management*. 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*(3), 914–927. https://doi.org/10.1037/0096-1523.23.3.914
Earle, F. S., & Myers, E. B. (2015). Overnight consolidation promotes generalization across talkers in the identification of nonnative speech sounds. *The Journal of the Acoustical Society of America, 137*(1), EL91–EL97. https://doi.org/10.1121/1.4903918
Fenn, K. M., Nusbaum, H. C., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature, 425*(6958), 614–616. https://doi.org/10.1038/nature01951
Firke, S. (2019). *janitor: Simple tools for examining and cleaning dirty data*. R Package Version 1.2.0. https://CRAN.R-project.org/package=janitor
Green, T., Rosen, S., Faulkner, A., & Paterson, R. (2013). Adaptation to spectrally-rotated speech. *The Journal of the Acoustical Society of America, 134*(2), 1369–1377. https://doi.org/10.1121/1.4812759
Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review, 105*(4), 611–633. https://doi.org/10.1037/0033-295X.105.4.611-633
Hornsby, B. W. Y., Naylor, G., & Bess, F. H. (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear and Hearing, 37*(Suppl. 1), 136S–144S. https://doi.org/10.1097/AUD.0000000000000289
Hunter, C. R., & Pisoni, D. B. (2018). Extrinsic cognitive load impairs spoken word recognition in high-and low-predictability sentences. *Ear and Hearing, 39*(2), 378–389. https://doi.org/10.1097/AUD.0000000000000493
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., Fedorova, V., & Fernandez, G. (2017a). Cognitive factors as predictors of accented speech perception for younger and older adults. *The Journal of the Acoustical Society of America, 141*(6), 4652–4659. https://doi.org/10.1121/1.4986930
Ingvalson, E. M., Lansford, K. L., Fedorova, V., & Fernandez, G. (2017b). 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. (2015). Familiarization effects on consonant intelligibility in dysarthric speech. *Folia Phoniatrica et Logopaedica, 67*(5), 245–252. https://doi.org/10.1159/000444255
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
Lametti, D. R., Nasir, S. M., & Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *Journal of Neuroscience, 32*(27), 9351–9358. https://doi.org/10.1523/JNEUROSCI.0404-12.2012
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 Science, 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, 25*(2), 233–239. https://doi.org/10.1044/2015_AJSLP-15-0059
Lansford, K. L., & Liss, J. M. (2014a). Vowel acoustics in dysarthria: Speech disorder diagnosis and classification. *Journal of Speech, Language, and Hearing Research, 57*(1), 57–67. https://doi.org/10.1044/1092-4388(2013/12-0262)
Lansford, K. L., & Liss, J. M. (2014b). Vowel acoustics in dysarthria: Mapping to perception. *Journal of Speech, Language, and Hearing Research, 57*(1), 68–80. https://doi.org/10.1044/1092-4388(2013/12-0263)
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
Leach, L., & Samuel, A. G. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology, 55*(4), 306–353. https://doi.org/10.1016/j.cogpsych.2007.01.001
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*(6), 3022–3030. https://doi.org/10.1121/1.1515793
Liss, J. M., White, L., Mattys, S. L., Lansford, K., 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*(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*(1), 552–561. https://doi.org/10.1121/1.2931948
Masapollo, M., & Guenther, F. H. (2019). Engaging the articulators enhances perception of concordant visible speech movements. *Journal of Speech, Language, and Hearing Research, 62*(10), 3679–3688. https://doi.org/10.1044/2019_JSLHR-S-19-0167
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

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457. https://doi.org/10.1037/0033-295X.102.3.419

Peelle, J. E. (2018). Listening effort: How the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear and Hearing, 39*(2), 204–214. https://doi.org/10.1097/AUD.0000000000000494

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C., Naylor, G., Phillips, N., Richter, M., Rudner, M., Sommers, M., Tremblay, K., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The framework for understanding effortful listening (FUEL). *Ear and Hearing, 37,* 5S–27S. https://doi.org/10.1097/AUD.0000000000000312

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

RStudio Team. (2018). *RStudio: Integrated development for R*. RStudio. http://www.rstudio.com/

Robinson, T., & Hayes, A. (2019). *broom: Convert statistical analysis objects into tidy tibbles*. R Package Version 0.5.2. https://CRAN.R-project.org/package=broom

Sidaras, S. K., Alexander, J. E. D., & Nygaard, L. C. (2009). Perceptual learning of systematic variation in Spanish-accented speech. *The Journal of the Acoustical Society of America, 125*(5), 3306–3316. https://doi.org/10.1121/1.3101452

Spencer, K. A., & Dawson, M. (2019). Dysarthria profiles in adults with hereditary ataxia. *American Journal of Speech-Language Pathology, 28*(2S), 915–924. https://doi.org/10.1044/2018_AJSLP-MSC18-18-0114

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*(2), 230–238. https://doi.org/10.1044/2015_JSLHR-S-15-0271

Stipancic, K. L., Yunusova, Y., Berry, J. D., & Green, J. R. (2018). Minimally detectable change and minimal clinically important difference of a decline in sentence intelligibility and speaking rate for individuals with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research*, *61*(11), 2757–2771. https://doi.org/10.1044/2018_JSLHR-S-17-0366

Van Nuffelen, G., De Bodt, M., Vanderwegen, J., Van de Heyning, P., & Wuyts, F. (2010). Effect of rate control on speech production and intelligibility in dysarthria. *Folia Phoniatrica et Logopaedica, 62*(3), 110–119. https://doi.org/10.1159/000287209

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). Oxford University Press.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://doi.org/10.1007/978-3-319-24277-4_9

Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation*. R Package Version 0.8.3. https://CRAN.R-project.org/package=dplyr

Xie, X., Earle, F. S., & Myers, E. B. (2018). Sleep facilitates generalisation of accent adaptation to a new talker. *Language, Cognition and Neuroscience, 33*(2), 196–210. https://doi.org/10.1080/23273798.2017.1369551

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