JSLHR

Research Article

Effects of Lexical and Somatosensory Feedback on Long-Term Improvements in Intelligibility of Dysarthric Speech

Stephanie A. Borrie[a] and Martina C. M. Schäfer[b]

Purpose: Intelligibility improvements immediately following perceptual training with dysarthric speech using lexical feedback are comparable to those observed when training uses somatosensory feedback (Borrie & Schäfer, 2015). In this study, we investigated if these lexical and somatosensory guided improvements in listener intelligibility of dysarthric speech remain comparable and stable over the course of 1 month.
Method: Following an intelligibility pretest, 60 participants were trained with dysarthric speech stimuli under one of three conditions: lexical feedback, somatosensory feedback, or no training (control). Participants then completed a series of intelligibility posttests, which took place immediately (immediate posttest), 1 week (1-week posttest) following training, and 1 month (1-month posttest) following training.
Results: As per our previous study, intelligibility improvements at immediate posttest were equivalent between lexical and somatosensory feedback conditions. Condition differences, however, emerged over time. Improvements guided by lexical feedback deteriorated over the month whereas those guided by somatosensory feedback remained robust.
Conclusions: Somatosensory feedback, internally generated by vocal imitation, may be required to affect long-term perceptual gain in processing dysarthric speech. Findings are discussed in relation to underlying learning mechanisms and offer insight into how externally and internally generated feedback may differentially affect perceptual learning of disordered speech.

People can improve their ability to understand a speech signal that is initially difficult to recognize. This statement holds true for the degraded speech signal of dysarthria. In a series of empirically based behavioral studies, we have shown that familiarization, or perceptual training, with dysarthric speech yields improved signal processing in subsequent encounters (Borrie, McAuliffe, Liss, Kirk, et al., 2012; Borrie, McAuliffe, Liss, O'Beirne, & Anderson, 2012, 2013; Borrie & Schäfer, 2015; Lansford, Borrie, & Bystricky, 2016). Such improvement is termed *perceptual learning* and defined as "relatively long-lasting changes to an organism's perceptual system that improve its ability to respond to its environment and are caused by this environment" (Goldstone, 1998, p. 586). In essence, the speech perception system can fine-tune its performance through experience.

Externally generated lexical feedback has been evidenced to drive perceptual learning of ambiguous or impoverished speech. Davis, Johnsrude, Hervais-Adelman, Taylor, and McGettigan (2005) systematically manipulated feedback conditions during perceptual training and showed that higher-level lexical knowledge, not syntactic or sentence-level semantic information, was crucial for improved processing of an artificially distorted speech signal. We have found similar results with perceptual learning of dysarthric speech, demonstrating superior intelligibility improvements when the degraded speech signal is supplemented with orthographic transcriptions during training (Borrie, McAuliffe, Liss, Kirk, et al., 2012; Borrie, McAuliffe, Liss, O'Beirne, et al., 2012). These studies, along with a number of others (e.g., Francis, Nusbaum, & Fenn, 2007; Loebach, Pisoni, & Svirsky, 2010; Norris, McQueen, & Cutler, 2003) suggest that disambiguating the degraded auditory information may be key to the experience-dependent processes involved in fine-tuning the speech perception system. Vroomen, van Linden, de Gelder, and Bertelson (2007) described this type of learning

[a]Department of Communicative Disorders and Deaf Education, Utah State University, Logan
[b]New Zealand Institute of Language, Brain and Behaviour, University of Canterbury, Christchurch
Correspondence to Stephanie A. Borrie: stephanie.borrie@usu.edu

Editor: Krista Wilkinson
Associate Editor: Kate Bunton
Received October 29, 2016
Revision received January 20, 2017
Accepted February 9, 2017
https://doi.org/10.1044/2017_JSLHR-S-16-0411

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

as an error signal mechanism. The authors postulate that externally generated feedback (i.e., lexical knowledge regarding the spoken targets), allows the listener to detect discrepancies between the intended targets and their ambiguous realizations. This perceived discrepancy, or input mismatch, evokes an internal error signal, promoting mapping of the ambiguous realizations onto meaningful internal representations of speech (see also Guediche, Blumstein, Fiez, & Holt, 2014).

Another proposed driver of perceptual learning is internally generated somatosensory feedback. Adank, Hagoort, and Bekkering (2010) examined processing of an unfamiliar accent for individuals trained to attend to the accented speech under various feedback conditions and observed improved processing for those instructed to vocally imitate the accent. The authors hypothesized that the somatosensory feedback that arises from vocal imitation benefits speech processing by increasing neural activation in speech perception and production areas of the brain. In a follow-up study by the same lab, Adank, Rueschemeyer, and Bekkering (2013) used functional magnetic resonance imaging to validate their hypothesis, reporting increased activation in the left inferior frontal gyrus, supplementary motor area, and left superior temporal sulcus during processing of accented speech for individuals who imitated an unfamiliar accent during training. These neural changes were not evident in the individuals who simply repeated the spoken stimuli in their native accent. We recently demonstrated that imitation of dysarthric speech (spastic) can improve processing of the same degraded signal immediately following training (Borrie & Schäfer, 2015). Further, acoustic analysis of the imitation attempts revealed a significant positive correlation between imitation accuracy and intelligibility gain, validating the role of somatosensory feedback in perceptual learning of dysarthric speech.

The notion that somatosensory feedback can drive improved speech processing can be theoretically rooted in the directions into velocities of articulators (DIVA; Guenther, Hampson, & Johnson, 1998) model of speech production (see Borrie & Schäfer, 2015, for a detailed description of DIVA application to improved processing of dysarthric speech). In essence, although the DIVA model itself does not account for perceptual learning of speech, it acknowledges that trial-and-error attempts to match an acoustic target during vocal imitation link the auditory–perceptual and somatosensory reference frames of a speech sound map.[1] Thus, the DIVA model theory would suggest that vocal imitation guides perceptual learning of speech by linking the degraded acoustic-phonetic information with the somatosensory information required to produce the sound, thereby creating a new speech sound map (Golfinopoulos, Tourville, & Guenther, 2010). It could therefore be hypothesized that internally generated somatosensory feedback will induce greater perceptual learning of degraded speech than externally generated lexical feedback due to the creation of new speech sound maps.

It is interesting to note that our previous report of vocal imitation of dysarthric speech demonstrated that the intelligibility improvements immediately following perceptual training were equivalent across somatosensory and lexical feedback conditions. In light of the hypothesized fundamental differences in the learning mechanism underlying these two training conditions, this finding may appear somewhat unexpected and, on face value without further investigation, may refute the proposed hypothesis. Due to the significant theoretical and clinical implications, it is important that lexical and somatosensory feedback conditions are studied in more depth before conclusions are reached.

One behavioral method that affords further examination of feedback conditions in perceptual learning of degraded speech is investigation of training-induced performance gains over time. When applying the theoretical assumptions of the DIVA model along with neural findings of Adank et al. (2013) to the two distinct learning mechanisms in the context of longevity of perceptual learning, the proposed hypothesis requires refinement: Although training involving lexical feedback may initially serve to disambiguate the degraded acoustics, the associated perceptual adaptations will not persist long-term because novel neuropathway activation requires internally generated somatosensory feedback. To date, the longevity of intelligibility improvements following perceptual training with dysarthric speech, or any ambiguous or impoverished speech signal for that matter, has received limited attention in the literature. In an earlier study targeting the role of lexical feedback, we compared intelligibility scores immediately and 1 week posttraining for listeners trained with dysarthric speech stimuli under one of three conditions: lexical feedback, no feedback, or no training (Borrie, McAuliffe, Liss, Kirk, et al., 2012). The study found that intelligibility gains observed immediately following training for those trained with no feedback did not persist at 1 week follow-up. In opposition, some learning carryover was realized 1 week later for those trained with lexical feedback, although the perceptual benefit was significantly less than that reported immediately following training.

The purpose of this study was to investigate somatosensory and lexical learning mechanisms underlying improved processing of dysarthric speech by examining longevity of the intelligibility gain beyond 7 days. In order to do so, we first verified that our previous findings with spastic dysarthria were replicable and generalizable to a different form of dysarthria. In accordance, our initial research question addressed the following: Are intelligibility improvements immediately following training with ataxic dysarthria using lexical feedback equivalent to those immediately following training using somatosensory feedback? Given our previous findings, we hypothesized comparable immediate intelligibility gains for individuals receiving lexical or somatosensory feedback during perceptual training. Our central research question then addressed the

[1]The speech sound map, according to the DIVA model of speech production, refers to a set of cells, each one representing a different phoneme, high-frequency multiphonemic syllable, or word (Guenther, Ghosh, & Tourville, 2006).

following: Are immediate intelligibility benefits associated with training involving lexical or somatosensory feedback maintained 1 week and 1 month following training, and if so, are they comparable between the two feedback conditions? We hypothesized that condition differences would emerge over time, specifically that training involving somatosensory feedback, relative to lexical feedback, would result in more enduring perceptual benefit.

Method

Participants

Sixty young, healthy adults (53 women and 7 men) aged 19 to 33 years ($M = 22.11$, $SD = 2.42$) participated in the experiment. All participants were native speakers of American English and passed a pure-tone hearing screen at 20 dB HL for 500, 1000, 2000, and 4000 Hz in both ears. As per self-report, participants had no history of speech, language, or cognitive disorders and no prior experience with individuals with dysarthria. Participants were recruited from undergraduate classes at Utah State University, and institutional review board consent was obtained before the start of the experiment.

Speech Stimuli

Speech stimuli consisted of semantically plausible ($n = 40$) and semantically anomalous ($n = 120$) audio-recorded phrases elicited from a native speaker of American English with a clinical diagnosis of dysarthria secondary to cerebellar disease. The semantically plausible phrases (e.g., "the nasty weather caused severe flooding") ranged in length from four to eight words, containing between four and 12 syllables per phrase. These phrases served as the linguistically-rich training speech set in the experimental procedure. The semantically anomalous phrases (e.g., "amend estate approach") ranged in length from three to five words, containing exactly six syllables per phrase. These phrases, designed to reduce the influence of lexical cues on word recognition (Liss, Spitzer, Caviness, Adler, & Edwards, 1998), served as the testing speech sets. Four testing speech sets were created, each consisting of 30 novel phrases and balanced for number of words.

The 64-year-old male speaker who provided the speech stimuli presented with the cardinal features of an ataxic dysarthria of moderate severity as diagnosed by three independent speech-language pathologists with expertise in assessment and differential diagnosis of motor speech disorders. Perceptually, his speech was characterized by excess and equal stress (scanning speech), prolonged phonemes and intervals, monotone, monoloudness, and imprecise articulation with irregular articulatory breakdowns (Darley, Aronson, & Brown, 1975; Duffy, 2013).

Procedure

The 60 participants were randomly assigned to one of three conditions ($n = 20$): (a) control, (b) lexical feedback, and (c) somatosensory feedback. The experimental procedure, programmed as a listener perception application hosted on a secure university-based web server, was conducted in five phases: (a) pretest, (b) training, (c) immediate posttest, (d) 1-week posttest, and (e) 1-month posttest. The four testing phases were the same across the three conditions. However, the training phase differed, depending on condition assignment. Sessions were held in the Human Interaction Lab at Utah State University. Participants were seated in front of a computer preloaded with the listener perception application and informed that task-specific instructions would be delivered via the computer program. The speech stimuli used in the testing and training phases were presented binaurally at a comfortable listening level of 65 dB through sound-attenuating headphones (Sennheiser HD 650 PRO, Old Lyme, CT).

Testing Phases

Four testing phases were carried out: (a) prior to training (pretest), (b) immediately following training (immediate posttest), (c) 7 days following training (1-week posttest), and (d) between 28 and 31 days following training (1-month posttest). The procedure used for testing was identical to that used by in our previous studies (e.g., Borrie & Schäfer, 2015). Participants were informed that they would be presented with short phrases produced by someone with a speech disorder. They were told that the phrases all contained real English words but would not make sense. Following audio presentation of each phrase, participants were instructed to use the keyboard to type out what they thought was being said. Participants were encouraged to make a guess at any words they did not recognize and to use an *X* to represent any part of a phrase where guess could not be made. Once they had finished typing their response, participants were prompted to press the return key to move on to the next phrase. The presentation order of the test phrases was randomized for each participant.

Training Phase

Immediately following the pretest phase, participants in the two experimental feedback conditions engaged in the training phase, which used the same training speech set but included either lexical or somatosensory feedback. Participants were told that they would again hear a series of phrases produced by a person with a speech disorder but that no typing would be required. They were told that the phrases would make sense and that they would need to listen closely to each phrase and try to understand what was being said. Participants in the lexical feedback condition were also told that written subtitles of the spoken phrases would be provided on the computer screen and that they should use this information to help them understand the speech. Participants in the somatosensory feedback condition were also told to repeat each spoken phrase back, as best as possible, in the same manner (i.e., speaking rate, intonation, stress, and pronunciation) it was produced.

File Analysis

The total data set consists of 240 transcripts: 60 transcripts of the pretest and 60 transcripts of each posttest (i.e., immediate, 1-week, and 1-month). Transcripts were scored by one judge for a standard measure of speech intelligibility, percentage of words correct (PWC). Scoring of words correct followed the same procedures outlined in previous studies on perception of dysarthric speech (e.g., Borrie, McAuliffe, Liss, Kirk, et al., 2012; Liss, Spitzer, Caviness, & Adler, 2002). Words were scored as correct if they matched the intended target or differed by one tense (-ed) or plurality (-s). In addition, obvious misspellings, homophones, and substitutions between "a" and "the" were scored as correct. Four PWC scores, one for each test, were tabulated for each participant, reflecting intelligibility before and after training.

Reliability Analysis

Twenty percent of the total data set (transcripts) were randomly selected according to computer-generated random number lists and reanalyzed by the original judge (intrajudge) and a second trained judge (interjudge) to obtain reliability estimates for scoring of words correct. Reliability analysis revealed high agreement between the reanalyzed and original data with Pearson correlation r scores above 0.971.

Results

Intelligibility data, expressed by the mean PWC scores, were collected for each participant at each test, and averaged for each condition. Figure 1 provides a summary of these results. To test the effect of test (pretest, immediate posttest, 1-week posttest, 1-month posttest) and condition (control, lexical, somatosensory) on intelligibility of dysarthric speech, PWC scores were analyzed using a two-factor mixed design analysis of variance. The within-subject factor was test and the between-subjects factor was condition.

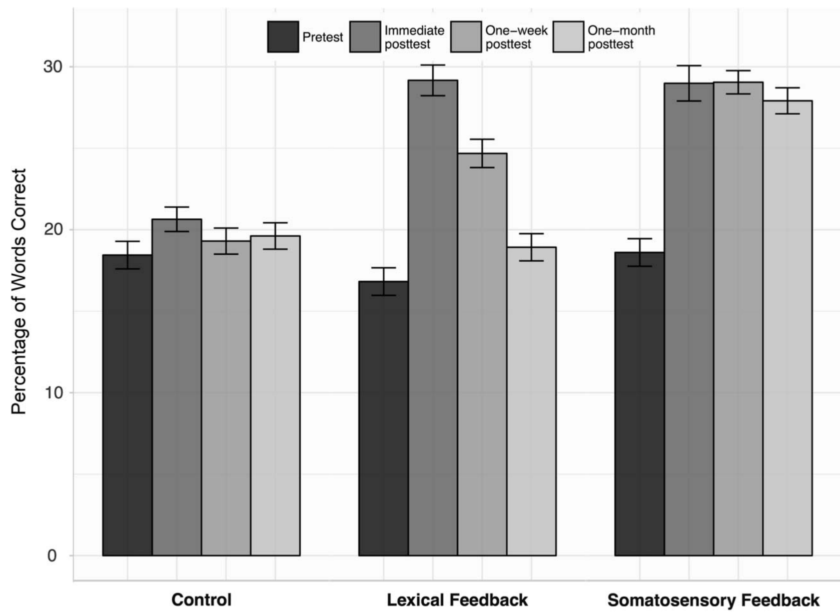
The analysis of variance showed a statistically significant interaction between test and condition on intelligibility of dysarthric speech, $F(6, 228) = 13.52$, $p < .001$, partial $\eta^2 = .26$. Simple main effects of condition were used to examine differences in PWC scores at each of the four levels of the within-subject factor test. In accordance, four one-way analysis of variance tests were carried out. For the pretest, no significant difference between conditions was noted. There was, however, a significant difference between conditions at immediate posttest, $F(2, 57) = 27.14$, $p < .001$, partial $\eta^2 = .52$; 1-week posttest, $F(2, 57) = 37.54$, $p < .001$, partial $\eta^2 = .57$; and 1-month posttest, $F(2, 57) = 37.76$, $p < .001$, partial $\eta^2 = .57$. Post hoc independent t tests with Bonferroni correction were conducted and are reported in Table 1. In essence, immediate posttest scores for the lexical and somatosensory conditions were both significantly greater than the control condition but were comparable to one another. At the 1-week posttest, scores for the lexical and somatosensory conditions were both significantly greater than the control condition, but those for the lexical condition were significantly less than the somatosensory condition. At the 1-month posttest, scores for the somatosensory condition were significantly greater than the control condition, but there was no difference between the lexical and control conditions.

Simple main effects of test were used to examine differences in PWC scores at each of the three levels of the between-subjects factor, condition. In accordance, three repeated-measures analysis of variance tests were carried out. For the control condition, no significant difference between tests was noted. There was, however, a significant difference between tests for lexical, $F(3, 57) = 42.19$, $p < .001$, partial $\eta^2 = .69$, and somatosensory conditions, $F(3, 57) = 39.58$, $p < .001$, partial $\eta^2 = .68$. Post hoc independent t tests with Bonferroni correction were conducted and are reported in Table 2. In essence, pairwise comparisons for the lexical condition showed that the immediate posttest scores were significantly greater than the pretest and other posttest scores. The 1-week posttest scores were also significantly greater than the pretest and 1-month posttest scores. However, there was no difference between the pretest and 1-month posttest scores. Pairwise comparisons for the somatosensory condition, on the other hand, showed that all posttest scores were significantly greater than the pretest scores, and there was no difference between any of these posttest scores.

Discussion

The purpose of this study was to examine longevity of intelligibility improvements following lexical and somatosensory guided perceptual training with ataxic dysarthria. Consistent with our hypotheses, the results showed that intelligibility improvements immediately following training were comparable across these two feedback conditions. This initial finding replicates the result observed in our previous study with spastic dysarthria, whereby training with orthographic transcriptions of the spoken targets was equally effective in enhancing immediate intelligibility gain as training that instructed vocal imitation of the degraded speech input (Borrie & Schäfer, 2015). A key contribution of the current study, however, is the ensuing finding that lexical and somatosensory guided intelligibility improvements did not remain comparable over time. As predicted, internally generated somatosensory feedback during training resulted in a more lasting perceptual gain than externally generated lexical feedback. To be specific, at both 1 week and 1 month following training, intelligibility scores for listeners trained with lexical feedback were significantly smaller than those for listeners trained with somatosensory feedback. Indeed, at the 1-month posttest, intelligibility scores for those in the lexical feedback condition had returned to levels equivalent to their pretest scores (and those of the control condition). On the other hand, intelligibility scores for those in the somatosensory feedback condition remained significantly higher

Figure 1. Average intelligibility, as measured by percentage of words correct, for listeners trained with dysarthric speech under different feedback conditions (n = 20) at four phases of testing. Control refers to no training. Error bars delineate ±1 *SEM*.



than their pretest scores. Further, the perceptual gain associated with somatosensory feedback did not deteriorate over time, suggesting an enduring change to the speech perception system. Findings are expanded upon below.

As summarized by Guediche et al. (2016), lexical feedback during perceptual training is thought to promote mapping of the ambiguous or impoverished acoustics onto existing representations of speech by evoking an internal error signal, arising from a discrepancy between the auditory input and externally provided lexical information. Here, the error signal arises from the mismatch between the input predicted by the orthographic transcriptions and the spoken realizations of the speaker with dysarthria. Immediately following training with lexical feedback, significant intelligibility improvements were observed. This initial finding is consistent with a number of previous studies and implicates fine-tuning of linguistic representations of speech (e.g., Borrie et al., 2013; Liss et al., 2002). Lexically guided intelligibility improvements, however, showed deterioration after 7 days, which is also consistent with an earlier study (Borrie, McAuliffe, Liss, Kirk, et al., 2012). Novel here is that lexically guided intelligibility improvements were no longer evident after 1 month. Taken together, the intelligibility data suggest that lexical feedback may not be sufficient to facilitate long-term change to the perceptual system for processing dysarthric speech. In light of Goldstone's (1998) definition of perceptual learning wherein "…relatively long-lasting changes to an organism's perceptual system…" (p. 586) are stipulated, the immediate intelligibility gain observed following training with lexical feedback may not represent true perceptual learning. Instead, the perceptual improvements may reflect a temporary fine-tuning of existing representations that, without any additional speech input, eventually return to canonical representations of linguistic knowledge.

Table 1. Results from the post hoc analyses of the simple main effects of condition at each level of posttest.

Comparison	*t* statistic	*p* value	Cohen's *d*
Immediate posttest			
Control versus lexical	6.45[a]	< .001	2.23
Control versus somatosensory	6.31[a]	< .001	2.01
Lexical versus somatosensory	0.14	1.000	0.04
1-week posttest			
Control versus lexical	4.77[a]	< .001	1.44
Control versus somatosensory	8.65[a]	< .001	2.88
Lexical versus somatosensory	3.88[a]	< .001	1.23
1-month posttest			
Control versus lexical	0.59	1.000	0.18
Control versus somatosensory	7.21[a]	< .001	2.31
Lexical versus somatosensory	7.81[a]	< .001	2.45

[a]Significant difference at $p < .001$.

The intelligibility improvement following training with somatosensory feedback, on the other hand, showed no signs of deterioration over the course of a month. Although immediate somatosensory guided perceptual benefits have been reported (Adank et al., 2010, 2013; Borrie & Schäfer, 2015), no previous studies have examined the longevity of such benefits. The results of the present study afford preliminary behavioral evidence for the hypothesis

Table 2. Results from the post hoc analyses of the simple main effects of test at lexical and somatosensory feedback conditions.

Comparison	*t* statistic	*p* value	Cohen's *d*
Lexical feedback condition			
Pretest versus immediate posttest	8.89[a]	< .001	3.09
Pretest versus 1-week posttest	6.21[a]	< .001	2.05
Pretest versus 1-month posttest	2.17	.260	0.56
Immediate posttest versus 1-week posttest	3.62[a]	.010	1.12
Immediate posttest versus 1-month posttest	7.73[a]	< .001	2.58
1-week posttest versus 1-month posttest	5.31[a]	< .001	1.51
Somatosensory feedback condition			
Pretest versus immediate posttest	9.64[a]	< .001	3.09
Pretest versus 1-week posttest	10.36[a]	< .001	2.99
Pretest versus 1-month posttest	8.14[a]	< .001	2.53
Immediate posttest versus 1-week posttest	0.05	1.000	0.02
Immediate posttest versus 1-month posttest	0.82	1.000	0.25
1-week posttest versus 1-month posttest	1.24	1.000	0.33

[a]Significant difference at $p < .02$.

that somatosensory feedback generates dysarthria-specific mental representations and that these representations remain active, accurate, and readily available for processing the degraded speech over the course of a month. Theoretical assumptions on the basis of the DIVA model and previous neuroimaging studies (Adank et al., 2013) would suggest that somatosensory feedback is required for lasting perceptual change. To be specific, vocal imitation is thought to link auditory and somatosensory reference frames for the degraded acoustics, thereby creating new dysarthria-specific speech sound maps and possibly novel neuropathway activations. It is assumed that these neural changes then improve processing of the degraded speech signal, likely due to the activation of perceptual and motor areas of the brain that reflect the auditory and somatosensory information of the initial teaching signal. The greatest benefit hereby, if neural pathways are indeed created and activated when listening to dysarthric speech in subsequent encounters, is the longevity of intelligibility improvements. Although neural data are needed to confirm these speculations regarding the underlying learning mechanisms, the present data support the idea that internally generated somatosensory feedback is required to effect lasting change to the speech perception system.

A concern that may arise with framing the current findings in DIVA theory is if improved processing of dysarthric speech comes at a processing cost for nondysarthric speech. This is not the case. Establishing a dysarthria-specific speech sound map is neither assumed to override nor temporarily adapt existing sound maps to include the degraded information. Rather, listeners treat the degraded speech as a new teaching signal and establish new mental representations, which represent a separate entity within the neural network. These speech sound maps then coexist and are activated—depending on the listening context—at the 1-week and 1-month posttests in the present study. The idea that specific speech sound maps are required for long-term processing of dysarthric speech is consistent with recent perceptual theories that suggest faster and more robust speech processing with separate generative models for an individual speaker or a group of speakers with a similar way of talking (Kleinschmidt & Jaeger, 2015).

How long a generative model for a specific speaker or speaker group is maintained and how accurately it will be maintained without additional input to validate the distributions of the belief system, however, are questions that have not been addressed in this theory. Similarly for the DIVA model, how long the information of a newly created speech sound map would remain easily accessible or valuable to maintain in the context of competing resources, given that it will likely not be required in everyday life, have been questions of unknown territory until now. In speech sound acquisition, the purpose of imitation is in the long-term use of a specific language repertoire. This may, for example, explain why children more easily learn new languages; universal speech sound maps acquired during the initial babbling phases may still be active to some degree despite not having been used for an extended period of time. The current results provide initial evidence that the generative model for or the speech sound map of a speaker with ataxic dysarthria remains active, accurate, and readily available for at least a month without further input or training. Further research is required to study the persistence of these learning mechanisms at 3 and 6 months.

In this study, we used speech stimuli produced by a single speaker whose speech features exemplified the cardinal features of ataxic dysarthria. As per a number of our previous studies, this was intentionally done for a high level of experimental control. Although the potential limitation of using a single speaker should be raised, it is perhaps more important to acknowledge that the results of this study may be limited to perceptual learning of dysarthria of classic ataxic presentation. However, we have previously reported comparable lexical and somatosensory guided immediate intelligibility improvements for spastic dysarthria (Borrie & Schäfer, 2015) and deterioration of lexically guided intelligibility gain for hypokinetic dysarthria (Borrie, McAuliffe, Liss, Kirk, et al., 2012). This suggests that the present results may also generalize to other types of dysarthria.

Another plausible limitation is that the results of this study are specific to perceptual learning of dysarthric speech of moderate severity. In other words, the value of lexical and somatosensory feedback may be differentially influenced by the severity of the speech degradation. Lexical feedback, for example, may afford less immediate perceptual gain when the intelligibility deficit is mild given the role it plays in disambiguating the degraded acoustics. Severe intelligibility impairment, on the other hand, may require combined somatosensory and lexical feedback to successfully overcome the increased challenge of imitating and disambiguating the impoverished acoustic realizations. Examining the effects of externally provided and internally generated feedback on perceptual learning of dysarthria of ranging severities, therefore, provides an important future direction for this work.

Although largely theoretical, the present study offers some valuable clinical considerations. There is a large body of evidence suggesting that listener-targeted perceptual training paradigms may be a viable clinical tool for the remediation of intelligibility disorders associated with dysarthria (see Lansford et al., 2016, for a summary). Here, we demonstrate that the inclusion of somatosensory feedback may be critical to the development of such training tools, ensuring long-term perceptual change to listener processing of dysarthric speech. Future studies, however, are needed to investigate additional training factors, such as amount and frequency of feedback and generalization of learning, before conclusions regarding best practices for listener-targeted perceptual training paradigms can be made.

Conclusion

In sum, the current study offers new insight into the influence of lexical and somatosensory feedback on perceptual learning of dysarthric speech. Here, we see that although both feedback conditions support immediate intelligibility improvements—a finding previously noted in the literature—somatosensory feedback is required to facilitate an enduring perceptual change. This finding can be theoretically framed within the DIVA model of speech production, which submits that both auditory and somatosensory information is required for successful speech sound acquisition and thereby indirectly also perceptual learning. The present study, therefore, highlights a critical role for internally generated somatosensory feedback in perceptual learning of dysarthric speech.

Acknowledgments

We gratefully acknowledge Michelle Parker and Ashley Hurst, research assistants in the Human Interaction Lab at Utah State University, for assistance with data collection and transcript analysis. We also acknowledge Tyson Barret from the Utah State University Statistical Consulting Studio for providing statistical input.

References

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science, 21,* 1903–1909.

Adank, P., Rueschemeyer, S. A., & Bekkering, H. (2013). The role of accent imitation in sensorimotor integration during processing of intelligible speech. *Frontiers in Human Neuroscience, 7,* 634. https://doi.org/10.3389/fnhum.2013.00634

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,* 1039–1055.

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

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

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

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

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

Duffy, J. R. (2013). *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management* (3rd ed.) St. Louis, MO: Elsevier Mosby.

Francis, A. L., Nusbaum, H. C., & Fenn, K. (2007). Effects of training on the acoustic phonetic representation of synthetic speech. *Journal of Speech, Language, and Hearing Research, 50,* 1445–1465.

Goldstone, R. (1998). Perceptual learning. *Annual Review of Psychology, 49,* 585–612.

Golfinopoulos, E., Tourville, J. A., & Guenther, F. H. (2010). The integration of large scale neural network modeling and functional brain imaging in speech motor control. *Neuroimage, 52,* 862–874.

Guediche, S., Blumstein, S. E., Fiez, J. A., & Holt, L. L. (2014). Speech perception under adverse conditions: Insights from behavioral, computational, computational, and neuroscience research. *Frontiers in Systems Neuroscience, 7,* 126.

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

Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language, 96,* 280–301.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review, 105,* 611–633.

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,* 148–203.

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,* 233–239.

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

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

Loebach, J. L., Pisoni, D. B., & Svirsky, M. A. (2010). Effects of semantic context and feedback on perceptual learning of speech processed through an acoustic simulation of a cochlear implant. *Journal of Experimental Psychology, 36,* 224–234.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47,* 204–238.

Vroomen, J., van Linden, S., de Gelder, B., & Bertelson, P. (2007). Visual recalibration and selective adaptation in auditory-visual speech perception: Contrasting build-up courses. *Neuropsychologia, 45,* 572–577.