Adaptive learning response to auditory perturbation of voice quality feedback.

Kari Urberg Carlson, Benjamin Munson, and Peter Watson

Department of Speech-Language-Hearing Sciences, University of Minnesota

Muscle Tension Dysphonia (MTD) is a common voice disorder whose cause is not well understood.

MTD is a disorder in which a patient is hoarse in the absence of (*primary MTD*) or disproportionate to (*secondary MTD*) any organic disorder of the larynx (Verdolini, Rosen & Branson, 2006). MTD is among the more common voice disorders, with primary MTD accounting for approximately 10-40% of the caseload of a typical voice center (Roy, 2003). Several factors have been proposed as causes of MTD: psychological/personality factors, vocal abuse/misuse, and compensation for an organic voice disorder that may or may not have resolved (Van Houtte, Van Lierde & Claeys, 2011). It is likely that all of these factors can cause MTD, but as of yet no mechanism for how this occurs has been proposed. The current study proposes adaptive learning as a mechanism for how compensation for an organic voice disorder can lead to MTD.

Adaptive learning has been demonstrated for several different speech behaviors.

Adaptive learning models can potentially explain MTD. Each time someone makes a movement, the adaptive learning system compares what it expects to happen to what actually happened. If there is a mismatch, it makes changes to the instructions that are sent to the muscles the next time that movement is made. For example, if a person puts on heavy roller skates, their brain notices that their feet are moving slower than expected, and uses more force to lift the legs. When the roller skates are removed, the feet move too fast until the brain adjusts again to use less force.

Adaptive learning has been shown to occur in a number of different types of speech behaviors. For example, when the pitch of a person's voice is artificially changed so that it sounds slightly higher or lower than it actually is, they change the pitch they are producing in the opposite direction to compensate (Larson, 1998). Other studies have shown that people's articulation of vowels (Houde and Jordan, 2002) and consonants like /s/ (Schiller, Sato, Gracco, & Baum, 2009) changes when their feedback is altered.

Adaptive learning of voice quality behaviors has not previously been examined.

Although adaptive learning has been demonstrated for perturbation of pitch and vowel height and backness, no experiments so far have looked at perturbation of voice quality- whether the voice is clear, hoarse, breathy, or strained. Voice quality differs from vowel height and backness in that it never carries lexical information in English, only indexical information such as gender, sexuality, emotion, and state of health. In studies of perturbation of pitch, in experiments where pitch carried only indexical information, participants usually compensated but sometimes changed their pitch in the same direction as the perturbation (a *following response,* e.g. Burnet, Senner and Larson, 1997). In studies where pitch carried lexical information, such as tone in Mandarin (Jones and Munhall, 2002) or word stress (Natke and Kalveram, 2001), the direction of change was always compensatory. Because voice quality carries only indexical information, we would predict that perturbation of voice quality feedback could cause either a compensatory or a following response.

Acknowledgements

This research was supported by student travel and dissertation funds from the College of Liberal Arts at the University of Minnesota, and by a Doctoral Dissertation Fellowship at the University of Minnesota. The authors would like to thank Jayanthi Sasisekaran, Juergen Konczak, Sarah Schellinger and Gerald Burke for their contributions to this research.

Adaptive learning of voice quality behaviors may explain the development of MTD.

We conducted the current study to investigate whether adaptive learning processes explain how MTD occurs. The inflammation associated with laryngitis makes it difficult or impossible for a person to produce a normal voice quality. The inflamed vocal folds don't always close properly, making the voice sound breathier. Some people respond to laryngitis by resting their voices, some respond by simply producing a breathy voice, while others respond by trying to counter the breathiness by contracting the muscles in the larynx so that the vocal folds close more completely. The resulting voice quality is especially strained and hoarse. Our research proposes that when changes to the larynx make the voice sound different—like those that occur during laryngitis— the brain changes the instructions that it sends to the muscles, resulting in a hoarse-sounding voice. In some speakers, this adaptive process becomes habituated, resulting in the maintenance of a strained, hoarse voice after laryngitis has resolved.

The current study examines looks for evidence of adaptive learning of voice quality behaviors.

Sixteen participants with no history of speech, language or hearing disorders participated in this study. Each participant first recorded the syllable /ha/. A steady-state section of the vowel was inverse-filtered and re-synthesized with different levels of signal-to-noise ratio (SNR) using software developed by the Bureau of Glottal Affairs at UCLA (Kreiman, Gerratt & Antoñanzas-Barroso, 2006). SNR ranged from -20 to -11 dB. The purpose of the re-synthesis was to make the participants' own productions sound as if the were breathier than they really were. For the adaptive learning experiment, participants repeated the syllable 'ha' as the word appeared on the screen for .5 seconds once per second. While they spoke, they heard noise over headphones. The type of noise varied according to the experimental condition:

Table 1:Experimental conditions

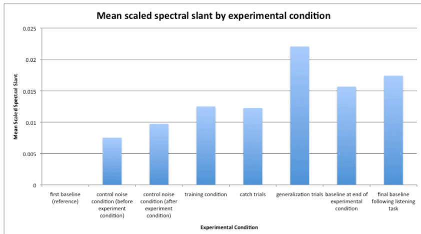
Condition	Order of presentation	What participants said	What participants heard	Purpose
Baseline	1	/ha/	Unmodified production	Establish baseline voice quality
Control	2 or 8	/ha/	Speech babble	Control for vocal fatigue and effect of added noise
Baseline	3 or 9	/ha/	Unmodified production	Establish baseline voice quality
Ramp-up	4 or 2	/ha/	Modified SNR of -20, -17, -14 dB	Gradually increase perceived breathiness
Training	5 or 3	/ha/	Modified SNR of -11 dB	Train motor behavior with perturbed feedback
Catch trials	6 or 4	/ha/	Unmodified production	Test for adaptation
Generalization trials	7 or 5	/he/	Unmodified production	Test for generalization to other utterances
Ramp-down	8 or 6	/ha/	Modified SNR of -11, -14, -17, -20 dB	Gradually decrease perceived breathiness
Baseline	9 or 7	/ha/	Unmodified production	Check for return to baseline behavior
Listening task				
Baseline	10	/ha/	Unmodified production	Check for return to baseli

The participants' speech was recorded on a Marantz CD-R 300 CD recorder. Using Praat (Boersma & Weenink, 2011), the beginning and end of each production was marked and coded according to experimental condition. The spectral slant (H2-H1), which has been found to correlate with the open quotient of the glottal vibration cycle (Holmberg, et al., 1995) was measured at the midpoint of each vowel. Because each participant's spectral slant was on a different scale based on their fundamental frequency and direction of change, the values of spectral slant were scaled. The scaled score measures the absolute value of change in spectral slant for each condition, using the first baseline condition as a reference. according to the following formula where *s* refers to the spectral slant:

$$s_{scaled} = \left| \frac{(\bar{s}_{condition} - \bar{s}_{baseline})}{\bar{F}_{0condition}} \right|$$

Participants' voice quality changed more in response to simulated breathiness than to speech babble.

Preliminary results indicate that the participants' voices had values of spectral slant that were more different from baseline during the training phase than during the control phase. The values were also more extreme during the catch trials, which indicates that adaptive learning may have occurred. The high scaled value of the generalization trials is likely an artifact that indicates that this acoustic measure is not stable across different vowels. Some of the data were not included in the analysis because the value of F0 calculated by Praat was inaccurate, compromising the accuracy of the spectral slant.



Future work will look for correlations with perceptual discrimination, and evidence of adaptive learning in clinical populations.

A second experiment examined the participants' ability to discriminate levels of hoarseness. Villacorta, Perkell and Guenther (2007) showed that participants who had better auditory discrimination of F1 compensated more for perturbed auditory feedback of F1 than those with poorer discrimination. This experiment will examine whether participants with better auditory discrimination of breathiness compensate more for perceived breathiness than those with poorer discrimination.

Patients with cerebellar damage have been shown not to adapt to other types of sensorimotor disruption. We predict that patients with ataxic dysarthria, which usually indicates cerebellar damage, would not adapt to auditory feedback of voice quality, and would not develop MTD subsequent to the cerebellar damage. Participants with a history of speech sound disorders may also be less likely to adapt to perturbed auditory feedback.

Clinical Implications

These preliminary results suggest that voice rest during acute episodes of laryngitis may be a valuable tool to prevent the onset of MTD, not because it prevents injury to the vocal folds, but because it avoids training the sensorimotor system in the presence of altered feedback.

References

Boersma, P. & Weenink, D. (2011). Praat: doing phonetics by computer [Computer program]. Version 5.2.19, retrieved 3 March, 2011 from http://www.praat.org/

Holmberg, E., Hillman, R., Perkell, J., Guiod, P., & Goldman S. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech and Hearing Research* **38**, 1212-1223.

Houde, J. F., & Jordan, M. I. (2002). Sensorimotor adaptation of speech I: Compensation and adaptation. *Journal of Speech, Language, and Hearing Research: JSLHR*, *45*(2), 295-310.

Jones, J. A., & Munhall, K. G. (2002). The role of auditory feedback during phonation: studies of Mandarin tone production. *Journal of Phonetics, 30*, 303–320.

Kreiman, J., Gerratt, B., & Antoñanzas-Barroso, N. (2006). *Analysis and Synthesis of Pathological Voice Quality.* Retrieved from http://www.surgery.medsch.ucla.edu/glottalaffairs/download.htm

Larson, C. R. (1998). Cross-modality influences in speech motor control: the use of pitch shifting for the study of F0 control. *J Commun Disord, 31*(6), 489-502; quiz 502-483; 553.

Natke, U., & Kalveram, K. T. (2001). Effects of frequency-shifted auditory feedback on fundamental frequency of long stressed and unstressed syllables. *J Speech Lang Hear Res, 44*(3), 577-584.

Roy. N., Bless, D.M. & Heisey, D. (2000). Personality and voice disorders: a multitrait-multidisorder analysis. *J Voice, 14,* 521–548.

Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America, 125*(2), 1103-1113.

Van Houtte, E., Van Lierde, K., & Claeys, S. Pathophysiology and treatment of muscle tension dysphonia: a review of the current knowledge. *J Voice, 25*(2), 202-207.

Verdolini K, Rosen C, Branski R (Eds). Classification manual for voice disorders-I. Mahwah, NJ: Lawrence Erlbaum Associates, 2006:249-252.

Villacorta, V. M., Perkell, J. S., & Guenther, F. H. (2007). Sensorimotor adaptation to feedback perturbations of vowel acoustics and its relation to perception. *J Acoust Soc Am, 122*(4), 2306-2319. .

.