Are Instructions to Manipulate Specific Parameters of Laryngeal Function Associated with Auditory-Perceptual Ratings of Voice Quality in Nondisordered Speakers?

Catherine J. Madill, Christine Sheard, and Robert Heard, *Lidcombe, New South Wales, Australia*

Summary: Objectives/Hypothesis. This preliminary study investigated whether auditory-perceptual judgments of voice quality by experienced speech language pathologists were associated with instructions given to speakers to manipulate specific laryngeal postures.

Study Design. Experimental, within-subject design.

Methods. Nine speakers were instructed and trained to manipulate three vocal parameters implicated in functional voice disorders—false vocal fold constriction, vocal fold mass, and larynx height—while reading a standard passage. Experienced judges rated these standard passages in terms of the widely used perceptual voice qualities breathiness, roughness, vocal strain, glottal fry, tone onset, tone color, loudness, and pitch.

Results. Between-subject factorial analysis of variance, controlling for judge unreliability, revealed some evidence that perceptual judgments are strongly associated with underlying laryngeal muscle activity that can then serve clinical planning of goals for intervention. Eta^2 effect sizes were large for all dependent measures, ranging from 0.39 for pitch to 0.77 for strained.

Conclusions. Although these results are encouraging, they were obtained under ideal conditions. Further research is warranted.

Key Words: Auditory-perceptual features–Laryngeal function–False vocal fold–True vocal fold mass–Larynx height.

INTRODUCTION

Auditory-perceptual judgments in the clinical setting

The validity and reliability of auditory-perceptual measures to evaluate vocal function and therapy outcomes in the clinic have been a focus of research for many years. An extensive range of variables has been identified as influencing the reliability of auditory-perceptual rating of the vocal signal including variability of terms, choice of speech sample, rating of normal versus pathological voices, listener experience and training, use of anchor stimuli and comparative ratings, scale type, reliability of judging specific auditory-perceptual features versus global voice quality,[1,2] and saliency of auditory-perceptual features.[3] There are many challenges to the reliability of auditory-perceptual judgments including instability of listener's internal standards, difficulties isolating individual attributes in voices, scale resolution, and the magnitude of the attribute being measured.[4,5] Despite the range of variables that can influence reliability of auditory-perceptual rating of the vocal signal, and the difficulty in establishing acceptable reliability in judgments, auditory-perceptual rating is still the most common form of clinical evaluation of the voice using informal strategies or rating tools such as the CAPE-V,[6] RBH,[7] or GRBAS[8] scales.[9]

Interpreting auditory-perceptual rating

Although perceptual judgments or ratings describe the nature of a voice as it is heard by the listener, interpreting these judgments has been difficult because the internal anatomy and function of the larynx and vocal tract are hard to specify objectively.[10] The perceptual consequences of changes in physical parameters and glottal vibratory patterns have been investigated in a few studies.[11] Some clinical outcome studies have described the provision of specific therapy techniques that target generalized movements, in which a range of co-occurring muscular movements are correlated with the resultant changes in perceived overall voice quality and other measures of vocal function.[12] Nevertheless, disordered vocal movement patterns across individuals are variable, as are their responses to therapy techniques.[13] Experimental research has investigated correlations between single isolated biomechanical muscular events and single auditory-perceptual features,[14] and simulation studies have also demonstrated the impact of changes in vocal tract shaping on auditory-perceptual judgments.[15,16] Although their aims are consistent with the need to correlate perceived features and laryngeal events, these studies rarely acknowledge or refer to the co-occurring movements of other muscular and biomechanical structures in the vocal tract.

The complexity of laryngeal functioning and the fact that specific laryngeal functions are unlikely to occur in isolation but rather, always in the context of other vocal tract functions, mean that ideal attempts to correlate isolated laryngeal movements with perceived voice features are likely to be inadequate in reflecting laryngeal reality. As a result, "perceptual evaluation will remain limited in its clinical application until we know more about its relationships to vocal tract function" (p. 52).[10] Consequently, this area needs further investigation.

Accepted for publication October 11, 2016.

From the Discipline of Speech Pathology, Faculty of Health Sciences, The University of Sydney, Lidcombe, New South Wales, Australia.

Address correspondence and reprint requests to Catherine J. Madill, Discipline of Speech Pathology, Faculty of Health Sciences, The University of Sydney, PO Box 170, Lidcombe, NSW 1825, Australia. E-mail: cate.madill@sydney.edu.au

Journal of Voice, Vol. ■■, No. ■■, pp. ■■-■■

0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.10.008

Auditory-perceptual correlates of specific laryngeal functions

Auditory-perceptual correlates of specific changes in laryngeal biomechanical function have been investigated in two ways. Firstly, studies have matched visual judgments or descriptions of clinically observed biomechanical functioning or manipulations of vocal function with generic voice quality descriptors such as "throaty,"[17] or "breathy," "flow," "neutral," and "pressed," without reference to the presence or interaction of other auditory-perceptual features.[18]

Other investigations have reported the outcome of specific, co-occurring, manipulated muscular parameters of the vocal tract on acoustic, but not perceptual measures.[19,20] Only one study has investigated the perceptual correlates of four biomechanically defined voice "modes."[21] Unfortunately, the perceptual rating system (including the voice quality labels and rating scales) was not described in that study, although the function parameters were well defined. Of those, three vocal function parameters are commonly described and treated in functional voice disorders: false vocal fold (FVF) activity, true vocal fold mass (TVFM), and larynx height (LH).

FVF activity

The functional and anatomical link between the action of the true vocal folds (TVFs), and the FVFs is well established.[22,23] Excessive medialization of the FVFs interferes with periodic vibration of the TVFs to make their vibration more irregular[1,24] and slower.[25] The resultant acoustic noise caused by aperiodic vibration of the TVFs (particularly in the context of increased supraglottic muscular activation) results in auditory perception of parameters such as hoarseness, roughness, throatiness, and strain, and may also contribute to perceptions of breathiness.[17,26] These aperiodic features have been correlated also with decreases in volume and fundamental frequency, and therefore pitch perception.[27,28]

Retraction or abduction of the FVFs in combination with manipulation of the TVFs has been documented in clinical studies[29,30] and has been recommended for the treatment of hyperadductive voice disorders such as vocal nodules.[31,32] It is described in the yawn-sigh technique,[33] confidential voice therapy,[34] and resonant voice therapy,[12] and produces voice quality described as being clear and resonant, without roughness, strain, or hoarseness.[35,36]

True vocal fold mass

The concept that the vibrating mass of the TVFs changes with length and tension is well documented.[28] The term "mass" here refers to the change in the vibrating vertical depth of the medial surface of the TVF as the "cover" of the TVF is stretched.[4,28] When the rate of vibration of the TVFs increases (and fundamental frequency is increased), then their vibrating mass decreases and the surface area of each vocal fold that contacts the opposing fold is also reduced.[4,8] Thus, TVFM is closely related to TVF closure.

Changes in TVFM are typically associated with notions of vocal register.[28,37,38] There is contention as to the number of and labels for different registers[39]; however, the vocal registers of vocal or glottal fry, modal voice, head voice, and falsetto have all been associated with differing vocal fold mass and different specific auditory-perceptual features.[28,38]

Vocal fry (also known as glottal fry) or pulse register is generally perceived as a series of localized pulses and gaps.[25] It results from slow, pulse-like vibrations of short, lax, massed TVFs[40] and medialization of the FVFs.[25] Modal voice has increased loudness and lower pitch than head voice and falsetto. It results from the periodic vibration of thickening and shortening the TVFs.[41] Head voice is perceived as higher in pitch, with a "lighter" sound and less volume than modal register.[28] Head voice results from periodic vibration of longer but tenser TVFs than used in the production of modal voice. There is less vibrating mass and less closure of the TVFs.[42,43] Falsetto is the highest in pitch of the vocal registers, and the one with the least volume. Breathiness is also reported in this register.[38,41] TVFM is reduced as the tension of the TVFs increases.[44] The nature of TVF closure, the state of TVFM, and vocal register have also been linked to types of tone onset.[45–47]

Larynx height

Changing the length of the vocal tract changes auditory-perceptions of resonant qualities in the voice.[28] Changes or manipulation of LH results in changes in perceptions of voice as being "darker" or "brighter" in quality, even though it might not have a lower or higher pitch.[28,48] The impact of changing vertical larynx position on medial compression of the TVFs has been described,[49,50] as has the resultant auditory-perceptual quality.[51] Specifically, a raised larynx position may increase supraglottic muscular tension and medial compression of the TVFs, and it has been suggested that lowering the larynx reduces the degree of medial compression of the TVFs due to the action of tracheal pull.[30,32,35,50]

Interaction of perceptual features

As described above, although some individual auditory-perceptual features have been associated with individual elements of vocal function, it is apparent that changing one biomechanical parameter of the larynx and vocal tract may also result in changing another element of laryngeal function. Such interactions may result in a variety of auditory-perceptual features that as yet are not extensively documented. Interactions of the TVFs and supraglottic and resonant chambers produce complex acoustic results. Co-occurring auditory-perceptual features also interact with one another, i.e. the presence of one auditory-perceptual feature can influence perception of a different auditory-perceptual feature. Listeners will perceive some co-occurring auditory-perceptual features as more salient than others.[52]

The saliency of perceived co-occurring vocal traits is demonstrated in the use of multidimensional scaling. For example, Bergan and Titze[53] found that perceptions of roughness were influenced by fundamental frequency (*fo*), such that a lower *fo* was perceived as rougher than higher *fo*s. Similarly, Bele[2] revealed that loudness influenced the reliability of listeners' ratings of other parameters such as "ringing" and vocal fry. These findings suggest that changing a biomechanical parameter of voice that results in the emergence of a dominant auditory-perceptual feature may cause a change in the perception of other features that are not as salient to the listener.

Despite researchers proposing that full differentiated control of laryngeal function is possible,[46] it is necessary to recognize