Perspectives of the ASHA Special Interest Groups
SIG 12, Vol. 1(Part 3), 2016, Copyright © 2016 American Speech-Language-Hearing Association

One's Voice: A Central Component of Personal Factors in Augmentative and Alternative Communication

Rupal Patel

Departments of Communication Sciences and Disorders and Computer and Information Science,
Northeastern University
Boston, MA
VocaliD Inc.
Belmont, MA

Travis T. Threats

Department of Communication Sciences and Disorders
Saint Louis University
St. Louis, MO

Disclosures
Financial: Rupal Patel is the founder of VocaliD Inc., which is referenced in the paper. Travis T. Threats has no relevant financial interests to disclose.
Nonfinancial: Rupal Patel has no relevant nonfinancial interests to disclose. Travis T. Threats served as the guest editor for this Part of *Perspectives*.

Abstract

Augmentative and alternative communication (AAC) devices have opened the gates to interaction for those with severe communication impairments. In the assessment and intervention, all components of the World Health Organization's International Classification of Functioning, Disability, and Health (ICF) should be addressed. However, an important Personal Factor to full integration has been largely ignored—that of one's voice. Each one of us has a unique voice that conveys our age, cultural background and personality—it's how people know and remember you. These affordances of the natural voice are not available to those who express themselves using AAC devices. A personalized digital voice brings the field of AAC to a closer realization of the social model of disability in which individuals are not defined by their disability and it is just one aspect of them. Access to a personalized voice uplifts the AAC user and provides an opportunity for social and emotional engagement that enhances quality of life.

Augmentative and alternative communication (AAC) methods and technologies afford those with severe communication disabilities a means to interact with the world around them. Given its roots in the medical model, the focus of AAC has been on transmission of information content. Similar to the medical model in general, assessment and intervention with persons who use AAC has focused on the Body Function and Body Structure components of the *International Classification of Functioning, Disability and Health (ICF).* Using this model, one focuses on the physical structural and limitation challenges of persons, as well as their basic cognitive and linguistic skills to determine the complexity and nature of the system chosen.

To the credit of researchers and clinicians in AAC, the Activity/Participation needs and certain Environmental and Personal Factors have been addressed. For example, for children their personal home and school needs are used to determine vocabulary for the AAC device. It is also considered best practice to consider persons' age and culture for choosing items. Clinicians involve key stakeholders such as family members and teachers on how to use or interact with persons who use AAC. Even with all the components of the ICF addressed at some level, there is still an overriding structure of the medical model.

While intervention with persons who need AAC devices has changed lives for the better, an important personal factor has not been addressed. That factor is the most personal of all—one's voice. Our voice, much like our face, is unique to us and an essential aspect of who we are. Recent advances in considering disability through a social lens imposes new challenges and opportunities for the field of AAC. In the social model of disability, it is not only important to focus on the individual with disability but also consider the role of society— people, policies, attitudes, and technologies—in fully integrating people with disability into all aspects of functioning within a society (Tregaskis, 2002). The social model is also an empowerment model. The person is not seen as only the recipient of help by those in medical and educational professions. They are to have a say in their own future, their own "voice." In the social model literature, one's "voice" is a metaphor for having a say in one's own destiny, the ability to have their say respected by not only medical and educational professionals, but by society in general. But one's voice is also a physical entity. The question is, can one really have a voice in the metaphorical sense if they do not even have their own unique physical voice?

Advancements in speech technology have focused on speech clarity and not uniqueness or authenticity. In the context of an AAC user, while a speech-generating device provides a means to express ideas and desires, the voice is that of the machine, not the user. There is limited connection in terms of the relevant Personal Factors of the person's and the machine's speech.

Physicist Stephen Hawking is perhaps the most well-known person who uses an AAC device with an American accented voice even though he is British. Although he has come to accept that voice as his own, there are many others who use the same voice. In fact, given the lack of voice options on AAC devices, it is not uncommon for several students in a classroom to use the exact same voice (Ridley, 2012). Even more disconcerting is that boys and girls alike may be using an adult male voice. Imagine the personal and cultural dissonance of using an adult male voice if you were a 14-year-old girl from rural Mississippi. Given the uniformity of voice options on devices, it is somewhat understandable that many families and users initially resist AAC devices, fearing that their child will never develop their own voice or will lose their voice to the machine. What is the impact on self and social identity for a young AAC user? Rather than reducing barriers and enhancing integration, generic sounding robotic voices with AAC users draw undue attention to the user's disability. Modern day technological advances and societal expectations demand the consideration of personal and social voice in AAC.

Each one of us has a unique voiceprint that conveys our personality, physical size, age, gender, race, intellectual ability, attractiveness, sexual orientation, and even personality (cf. Bachorowski & Owren, 1999; Collins, 2000; Feinberg, Jones, Little, Burt, & Perrett, 2005; Fellows, Remez, & Rubin, 1997; Fitch & Giedd, 1999; Hartman & Danhauer, 1976; Linville, 1998; Munson, McDonald, DeBoe, & White, 2006; Pierrehumbert, Bent, Munson, Bradlow, & Bailey, 2004; Smyth, Jacobs, & Rogers, 2003; Walton & Orlikoff, 1994; Zuckerman & Miyake, 1993). These affordances of the natural voice are not available to those who express themselves using even "state-of-the-art" AAC.

The challenge is not merely to match the person's AAC voice with someone of similar relevant Personal Factors. In the example of Hawking, one could simply find another British physicist and program in the needed words for his vocabulary. This would sound more "natural," but it does not take into account that each person has a right to be unique in their voice, to literally have their own voice.

Research shows that even those individuals with severe dysarthria are capable of producing vocalizations that can be mined to extract the essence of their vocal identity (Patel, 2002a, 2002b; Patel, 2003; Patel & Campellone, 2009). Even a single vowel contains enough "vocal DNA" to seed the voice personalization process. These findings are now driving technological innovations focused on creating personalized digital voices (Mills, Bunnell, & Patel, 2014). One such technology is being developed by the first author of this paper. VocaliD Inc. creates custom crafted digital voices that reflect the end user by combining the recipient's residual vocal abilities

with a database of recordings produced by a demographically and acoustically matched speaker from the company's Human Voicebank repository. People from around the world with different demographic and personal characteristics are being recruited for this effort. The result is a voice that sounds like the recipient in age, personality, and vocal identity, but is as clear and understandable as the matched speaker. Just as a blended wine has a signature flavor, a blended voice is uniquely crafted to convey the age, gender, and cultural and linguistic history of the recipient. These customized voices are not mere replicas of the recorded samples, they are original creations that fit and honor persons using them—bringing attention to their words and voice. A personalized voice empowers the recipient to engage in conversation and be heard in his or her own voice.

An individualized voice brings the field of AAC to a closer realization of the social model, that people with disabilities are not defined by their disability, it is just one aspect of them. Below are two quotes from the ICF that emphasize this point:

> It is important to stress, moreover, that ICF is not a classification of people at all. It is a classification of people's health characteristics within the context of their individual life situations and environmental impacts. It is the interaction of the health characteristics and the contextual factors that produces disability. This being so, individuals must not be reduced to, or characterized solely in terms of their impairments, activity limitations, or participation restrictions. (World Health Organization [WHO], 2001, p. 242)

> Every scientific tool can be misused and abused. It would be naıve to believe that a classification system such as the ICF will never be used in ways that are harmful to people … Individuals classed together under ICF may still differ in many ways. Laws and regulations that refer to ICF classifications should not assume more homogeneity than intended and should ensure that those whose levels of functioning are being classified are considered as individuals. (WHO, 2001, pp. 244–245)

Thus, to view AAC as only the successful transmission of information and basic transactions is to shortchange the person their "individual life situation," which includes their voice. Two children who use AAC devices could be in the same class in the same school with similar physical and cognitive-linguistic abilities. Using traditional criterion, they would use the same device. But both of these children still deserve to speak through their own distinct voice even if that is via an AAC device.

Within the construct of Activity/Participation, the match of the person to the voice might encourage these persons to interact more because of the naturalness of the speech produced. In communication there needs to be a willing sender and a willing receiver. The naturalness of the voice may affect both the speaker and the listener to more fully engage with each other. Communication is not simply for transactions (e.g., pass me the bread) but has aesthetic and social components. People simply enjoy talking with some people more than others for a variety of reasons, one of which can be speech and voice characteristics. It is possible that the mechanical voice, which has produced an environmental barrier, can be advanced to be an environmental facilitator.

In the annex concerning ethical use of the ICF, it states:

> Wherever possible, the person whose level of functioning is being classified (or the person's advocate) should have the opportunity to participate, and in particular to challenge or affirm the appropriateness of the category being used and the assessment assigned (WHO, 2001, p. 224)

Where in a typical AAC synthesized voice process does the person using the AAC device "participate" in making the voice that they will use in all of their verbal communications? Speech-language pathologists do have their clients give input to areas such as vocabulary words chosen and type of device used. There are often a limited number of voices to choose from for these

persons. Two people from the same region who are of the same background and age do not sound alike. Perhaps more than one blended voice could be offered and have the person choose which one they wish to use, which one is more them. Thus, at a basic ethical level the field needs to do more to ensure that they both literally and figuratively have a say and input in the creation of their voice. The voice is thus not made for them but with them. This sense that something in the speech produced by the device is really you is in itself empowering.

Implicit in the ICF framework is the concept of quality of life. However, as Cruice (2008) discusses, the use of the ICF does not ensure that the subjective aspects of persons' view of themselves, their quality of life, is adequately addressed. The ICF is not a quality of life scale because it often still relies on expert professional opinion to make the ratings. The domain most discussed in this paper, Personal Factors, is not coded in the ICF classification. Two future goals of the ICF framework are "establishing links with quality of life concepts and the measurement of subjective well-being" as well as "development of Personal Factors" components (WHO, 2001, p. 251).

In the American Speech-Language-Hearing Association's Scope of Practice for Speech-Language Pathology (ASHA, 2016) it states, "The overall objective of speech-language pathology services is to optimize individuals' abilities to communicate and to swallow, thereby improving quality of life." (ASHA, 2016, "Framework for Speech-Language Pathology Practice"). The implication here, as with health professions in general, is that this improvement automatically improves quality of life. This may be true, but it misses that there is sometimes more to life than being competent at doing something. There is no doubt that AAC devices improve quality of life by affording users a means to express their wants, needs, and thoughts. But, is it enough to simply improve functional communication? We as speech-language pathologists should do what we can to maximize quality of life. A personalized voice on an AAC device aspires toward that larger goal.

References

American Speech-Language-Hearing Association. (2016). Scope of Practice in Speech Language Pathology. Retrieved from http://www.asha.org/policy/SP2016-00343

Bachorowski, J., & Owren, M. (1999). Acoustic correlates of talker sex and individual talker identity are present in a short vowel segment produced in running speech. *Journal of the Acoustical Society of America, 106,* 1054–1063.

Collins, S. (2000). Men's voices and women's choices. *Animal Behavior, 60,* 773–780.

Cruice, M. (2008). The contribution and impact of the International Classification of Functioning, Disability and Health on quality of life in communication disorders. *International Journal of Speech-Language Pathology, 10,* 38–49.

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behavior, 69,* 561–568.

Fellows, J. M., Remez, R. E., & Rubin, P. E. (1997). Perceiving the sex and identity of a talker without natural vocal timbre. *Perception and Psychophysics, 59,* 839–849.

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America, 106,* 1511–1522.

Hartman, D., & Danhauer, J. (1976). Perceptual features of speech for males in four perceived age decades. *Journal of the Acoustical Society of America, 59,* 713–715.

Linville, S. (1998). Acoustic correlates of perceived versus actual sexual orientation in men's speech. *Folia phoniatrica et logopaedica, 50,* 35–48.

Mills, T., Bunnell, T. H., & Patel, R. (2014). Towards personalized speech synthesis for assistive communication. *Augmentative and Alternative Communication, 30,* 226–236.

Munson, B., McDonald, E. C., DeBoe, N. L., & White, A. R. (2006). The acoustic and perceptual bases of judgments of women and men's sexual orientation from read speech. *Journal of Phonetics, 34,* 202–240.

Patel, R. (2002a). Phonatory control in adults with cerebral palsy and severe dysarthria. *Augmentative and Alternative Communication, 18,* 2–10.

Patel, R. (2002b). Prosodic control in severe dysarthria: Preserved ability to mark the question-statement contrast. *Journal of Speech, Language, and Hearing Research, 45,* 858–870.

Patel, R. (2003). Acoustic characteristics of the question-statement contrast in severe dysarthria due to Cerebral Palsy. *Journal of Speech, Language, and Hearing Research, 46,* 1401–1415.

Patel, R., & Campellone, P. (2009). Acoustic and perceptual cues to contrastive stress in Dysarthria. *Journal of Speech Language and Hearing Research, 52,* 206–222.

Pierrehumbert, J., Bent, T., Munson, B., Bradlow, A. R., & Bailey, J. M. (2004). The influence of sexual orientation on vowel production. *Journal of the Acoustical Society of America, 116,* 1905–1908.

Ridley, L. [Lost Voice Guy]. (2012, December 13). *Voice by Choice comedy sketch by Lee Ridley, Lost Voice Guy* [Video File]. Retrieved from https://www.youtube.com/watch?v=CMm_XL3Ipbo

Smyth, R., Jacobs, G., & Rogers, H. (2003). Male voices and perceived sexual orientation: An experiment and theoretical approach. *Language and Society, 32,* 329–350.

Tregaskis, C. (2002). Social model theory: The story so far…. *Disability and Society, 17,* 457–470.

Walton, J., & Orlikoff, R. (1994). Speaker race identification from acoustic cues in the vocal signal. *Journal of Speech Language and Hearing Research, 37,* 738–745.

World Health Organization. (2001). *International Classification of Functioning, Disability, and Health*. Geneva, Switzerland: Author.

Zuckerman, M., & Miyake, K. (1993). The attractive voice: What makes it so? *Journal of Nonverbal Behaviour 17*(2), 119–135.

History:
Received April 13, 2016
Revised May 30, 2016
Accepted June 11, 2016
doi:10.1044/persp1.SIG12.94