Spring 5-5-2023

Talk Box in Music Therapy with Speech and Language Impairments Resulting From Tracheostomy: A Critical Review of the Literature

Jonathan Kang
jkang4@lesley.edu

Talk Box in Music Therapy with Speech and Language Impairments Resulting From Tracheostomy: A Critical Review of the Literature

Capstone Thesis

Lesley University

May 5th 2023

Jonathan Kang

Clinical Mental Health Counseling: Music Therapy

Dr. Rebecca Zarate

Abstract

Tracheostomy patients struggle with one of the most significant and essential components to the human identity, that is, communication. When a person's communication is obstructed, access to their intangible, inner world is also severed – leaving their thoughts, feelings, and memories, all of which are core to the human experience, unshared. Speech assistance for tracheostomy patients currently includes above cuff vocalization efforts using one-way speaking valves, augmentative and alternative communication methods such as speech generating devices and the electrolarynx. This thesis will introduce and discuss a musical instrument effect called the talk box as an integration of music therapy techniques that address speech and language deficiencies in tracheostomy patients. A talk box consists of a mini speaker amplifier with an input for instruments and an output where a plastic tube can be attached. The distal end of the plastic tube is placed in a speaker's mouth where they can use their articulators to form consonant and vowel shapes. The resulting effect amplifies the instrument's output into the speaker's mouth where words or phrases can be spoken or sung. Using an instrument with the talk box allows for a real-time adjustment of pitch, quality, loudness, timbre, and tempo. A large part of music therapy incorporates the use of music and instruments to help a patient with non-musical goals. The talk box could provide tracheostomy patients with a non-invasive and non-permanent approach to speaking and singing, as well as an opportunity to reclaim a missing part of their identity that cannot be replaced or replicated.

Keywords: tracheostomy; communication; speech generating devices; music therapy; talk box

Talkbox in Music Therapy with Speech and Language Impairments Resulting From Tracheostomy: A Literature Review

Introduction

Speech and language are essential components for human communication and interaction. It is what contributed to the survival of ancient civilizations and what allows us to progress into the future. However, when communication is impaired, it results in a disconnect that not only impacts the individual, but also the people and systems around them. When a person's communication is obstructed, access to their intangible, inner world is also severed – leaving their thoughts, feelings, and memories, all of which are core to the human experience, unshared. The American Speech-Language-Hearing Association (AHSA) shares that:

> "5% to 10% of Americans may have communication disorders. By the first grade, roughly 5% of children have noticeable speech disorders. 3 million+ Americans stutter. Nearly 7% of Americans have some form of language impairment. Approximately 1 million Americans suffer from aphasia" (ASHA, 2023)."

There are many factors that contribute to speech and language complications and while this thesis discusses different types and causes of speech and language disorders, it is focused on a demographic of people who suffer from speech and language disorders, specifically people who have tracheostomies.

Medical complications such as birth defects or cancer may result in patients getting surgery for clearing the airway and providing safe oxygen delivery to the lungs otherwise known as a tracheostomy. Many patients who receive the tracheostomy procedure may also be connected to a ventilator to help assist them in breathing. At a children's hospital in the Boston metro area, patients are on the pulmonary/respiratory rehabilitation unit where a vast majority of patients are tracheostomy patients. Due to the physical and invasive nature of a trach, patients are

not fully capable of breathing or producing sounds naturally. However, this does not prevent them from addressing their wants and needs. Patients are often able to communicate using non-verbal or non-vocal communication by using signs, gestures, and expressions.

With nonverbal populations, common functions of communication include more physical and gestural methods. Pointing, facial expressions, and eye-gazing are examples of pre-linguistic communication as discussed by van der Meer and Rispoli (2010). Supportive methods of accessibility such as communication boards as well as augmented and alternative communication (AAC) devices are commonly used to help people with communication disorders in self-expression. Blischak et al. (2003) explore the use of AAC, specifically the use of speech generating devices (SGD), that suggest possible reasons for promoting natural speech production.

Speech generating devices have been utilized in communication interventions to support natural speech for individuals with developmental disabilities and children with autism (Rispoli et al., 2010; Blischak et al., 2003; van der Meer & Rispoli, 2010). The work by Blischak et al. (2003) discusses many important qualities of communication such as rate of participation in conversation, increase in utterance length, reduced physical demands, reduction in pressure to speak, acoustic effects of SGD, pairing of graphics with spoken symbols, and the development of internal phonology. In other words, by using SGD and AAC, Blischak et al. discuss some of the measurable factors of communication, and furthermore encouraging research of natural speech production.

The present literature review will give an overview of topics such as communication, speech and language disorders, parts of the voice, and overlaps between music and speech. It also provides an outline of music therapy and speech pathology, discuss tracheostomies,

investigate current music therapy and medical interventions used for communication and speech, before turning the attention to alternative methods in assistive communication using devices.

This thesis will introduce and discuss a musical instrument effect called the Talkbox as an integration of music therapy techniques that address speech and language deficiencies in tracheostomy patients. The talkbox is an effect unit, that is most commonly an instrument effect pedal. It uses an external instrument such as a guitar or keyboard to produce vibrating sound waves that are processed, amplified, and passed through a plastic tube that goes into the mouth. The person then uses the shapes of the mouth to produce an output. A very recognizable use of this effect is from the opening seconds of "24k Magic" by Bruno Mars. Talkboxes have almost exclusively been used for musical performance and production – this thesis aims to encourage further research in using music technology for medical settings, specifically for adult patients with tracheostomies and offering them an alternative method of speech production and communication.

Method

Research for this thesis was primarily collected using Google Scholar on a web browser. The search criteria were mostly on literature since 2019 to ensure relevancy and recency but included publications ranging from 1964 to 2022. The reviewed literature consisted of peer-reviewed journal articles, book chapters, and informational web pages.

Search terms included string searches such as "music and speech", "music therapy and communication", and "music therapy and tracheostomy". Other search terms were: tracheostomy, talkbox music, talkbox, music therapy for tracheostomy patients, tracheostomy communication, what is communication, music therapy for speech disorders, speech generating

devices, intraoral electrolarynx, electrolaryngeal speech, connections with arts, social connection humans, and music therapy for socialization.

Literature references were organized using Zotero, a third-party software which embeds the web browser and allows the user to collect, cite, organize, and share resources. Furthermore, the literature was organized into themes of communication and speech disorders, tracheostomy patients, speech devices, and music therapy. The themes were researched individually and compiled as a whole.

Literature Review

Communication

Luhmann (1992) describes communication as the distinguished, intentional selection of "information, utterance, and understanding" (p. 252) and only by a combination of these three distinct components can it form communication. In essence, if information is the idea of thought and speech to be conveyed, understanding is the comprehension of the ideas, and utterance is how the information is delivered. Most commonly in a dialogue between two people, the utterance is observed as the spoken word of one or both individuals. In other words, the medium that is being used to deliver the information is using words and vocal speech. Adding to Luhmann's description of communication, I would modify the component of "utterance" to a more general idea that I will refer to as the "medium." In summary, communication is the sharing, exchanging, and understanding of information that takes place between two or more sources using a common medium.

Although the ability and capacity to communicate is natural with most people, there are suggestions that communication is learned rather than innate (Lapum et al., 2020). From birth,

and perhaps even before birth, a human is designed and wired to functionally communicate. Humans communicate their needs, wants, thoughts, and/or feelings with others, which result in socialization by forming relationships and sharing information.

There can be adverse effects of miscommunication and/or no communication. When one individual shares information but the other individual(s) do not understand what is being communicated, miscommunication occurs. Miscommunication can be represented as when one (or more) of the three components that Luhmann describes, information, [medium], and understanding is inadequate or missing. This can be at the fault of the person communicating their ideas, due to ambiguity or generalizations. This can also be at the fault of the person who is receiving the communication due to lack of knowledge, clarity, context, and/or insight. In the context of a romantic relationship between two people, miscommunication can often lead to hurt feelings, broken relationships, misunderstanding, and inefficiency in conversation. In the context of a therapeutic relationship, miscommunication may result in hurt or harm, inhibition to growth, inability to process, and ultimately resulting in severing the relationship.

No communication means no sharing of information. In other words, when a person does not or cannot communicate externally, their unique thoughts, feelings, and emotions are kept solely to the individual themselves. However, utilizing therapy, tools, and devices, people are able to convey their innermost selves externally.

Knowledge and Information

What is being communicated?

Swiss psychologist Jean Piaget's genetic epistemology establishes three types of knowledge, physical knowledge, logical-mathematical knowledge, and social knowledge. Piaget

describes physical knowledge as empirical, observed, and measurable such as the sky being blue. Logical-mathematical knowledge is abstract and unique to each individual, created by connections of experiences and invented relationships of information which cannot be heard, seen, or felt. Finally, social knowledge is described as man-made information that is specific to cultures and shared between them.

Expanding on Piaget's epistemology, knowledge and information can be acquired through observation, learning, and sharing. While knowledge that is gained through observation/learning is important, there are some types of information that can only be revealed by the one who knows that special knowledge. For example, an individual may observe another person based on their presented behavior and form a hypothesis that explains such behaviors. However, the observer will never know what the other person is thinking or feeling, with certainty, unless they share the information to the observer. When this scenario plays out in the real world, there is a significance to how the shared information is communicated.

How People Communicate

Communication involves many variables. Communication may occur between two or more individuals, between a human and a living non-human entity that is cognitively capable of understanding information, such as animals, or with the rise of machine learning, between humans and computers/artificial intelligence. What can we use and how are we able to create bonds and communicate with human and non-human entities?

Some ways of communicating include speech and language, non-verbal communication (eye movement, subtleties in face/mouth, body language, gestures), sign language, written

communication, and miscellaneous/other methods (subliminal messaging, morse code, braille, accessibility, arts/music).

Speech and language, categorized under verbal communication, are the most common methods of communication. The American Speech-Language-Hearing Association (n.d.) defines speech as "how we say words" and language as "the words we use to share ideas and get what we want" (ASHA, 2023). It is important to note that verbal speech may be among the most common methods of communication but is not exclusively the primary means of communication. This factors into more appropriately using the term "medium" to refer to the mode/method of communication which encompasses the "utterance" as described in Luhmann (1992) as a particular medium of communicating.

When humans communicate with other humans, a complex combination of speaking, gestures, movement, listening, and observing takes place. For example, a friend can be greeted with a high five, a smile, and saying, "Hello", where three simultaneous modes of communication are occurring: verbal, physical, and gestural. A similar scenario can take place, but a friend is greeted with a high five, a frown, and saying "hello". What happens when this same type of greeting is with a stranger rather than a friend? The scenario can look very different if the mode of communication is altered by the substitution, addition, or subtraction of different mediums.

Marshall McLuhan famously introduced the idea that the medium is the message. An excerpt from Understanding Media (1964):

> The instance of the electric light may prove illuminating in this connection. The electric light is pure information. It is a medium without a message, as it were, unless it is used to spell out some verbal ad or name. This fact, characteristic of all media, means that the "content" of any medium is always another medium. The content of writing is speech,

> just as the written word is the content of print, and print is the content of the telegraph. If it is asked, "What is the content of speech?," it is necessary to say, "It is an actual process of thought, which is in itself nonverbal." (p. 8)

In Mark Federman's discussion of McLuhan's theory, both the message and the medium are defined. According to Federman (2004), the message is "the change of scale or pace or pattern that a new invention or innovation introduces into human affairs" (p. 1) and the medium is "any extension of ourselves" (McLuhan, p. 7). Federman writes, "the message of the newscast are not the news stories themselves, but a change in the public attitude towards crime, or the creation of a climate of fear" (p. 2). Referring to the example of greeting a friend with a high five, a smile, and saying, "hello", the message is portraying hospitality and affection when analyzing the whole rather than the sum of its parts. In summary, McLuhan's worldview of a message encourages us to look past the face value of the content and to look at the bigger idea/picture of the content that is being shared.

Communication as a Multisensory - Multimodal Experience

Communication is an experience that is very complex and essential to the way people interact with the world around them. The process of sharing and/or exchanging information can be manifested in both verbal and non-verbal mediums. Communicating can look like movement and/or gestures, written messages, or most commonly, speech and language. There also are ways of communicating by using music and arts.

Holler and Levinson (2019) discuss the multimodal nature of communication, describing the variables of articulators and modalities. The articulators as described in Holler and Levinson generally refer to the physical components that humans use when producing speech, both verbal and sign language, which includes, "the tongue, lips, and mouth as well as the head, the face including the forehead and eyebrows, the upper and lower eyelids, the muscles around the nose,

cheeks, and mouth, the hands, arms, and shoulders, the upper torso, and, in principle, the lower torso, legs, and feet, although they tend to be less systematically used" (p. 640).

In addition to speech and language as a modality of communication where the primary articulators are the physical components of the human body, other "articulators" can be defined as the "medium". In sign language, the hands, arms, fingers, and facial expressions are the medium of communication. When looking at written forms of communication, the medium can be the pen or pencil, typewriter or keyboard, paper or digital screen and requires the receiver of communication to be able to read.

The music modality as a method of communication offers a multisensory experience. Music requires the person who is communicating to create something using their touch, hearing, and sight. When an individual communicates using music, many shared variables from both verbal and nonverbal communication occur simultaneously. Overlapping components of music and speech such as syntax, dynamics, and tempo will be discussed in a later section. How a person powerfully strikes a guitar chord or softly plays the keys of a piano communicates very different emotions and feelings despite playing the same musical composition.

Grossbach et al., (2011) discussed methods of communicating using "gestures, head nods, mouthing of words, writing, use of letter/picture boards and common words or phrases tailored to meet individualized patients' needs" (p. 46).

Understanding and Comprehension

The third component of Luhmann's definition of communication is understanding. When information is shared using a mode of communication, understanding the information is just as important, if not more, than the information and medium themselves. For example, when an

individual talks about a basketball game to a friend, what meaning would there be for the receiver if they know nothing about basketball? What does a signal sent using morse code, a message written in braille, or a string of characters of an unfamiliar language mean to someone who does not or cannot understand it? Without understanding, all that is perceived is a series of short and long beeps, raised bumps on a surface, or symbols that may resemble lines, shapes, or pictures.

When a patient with a tracheostomy tries to communicate by speaking, but their trach impairs their vocal cords, what is observed is mouth movements with minimal to no sounds. While trach patients could then find alternative methods of communication such as writing, gestures, sign language, or augmentative and alternative communication using speech devices, this requires the listener/receiver to understand a handful of things which include the articulators (or the medium), how the articulators/medium convey a message, and the message itself.

If a patient communicates using pen and paper, the listener must be able to read as well as decipher the patient's handwriting in order to understand the message. If sign language is used, the receiver must be able to discern what each of the signs mean. If speech generating devices are used, the listener must be able to distinguish between an urgent request, emotional thought, and narrative observation. Sometimes when communicating without speech, an interpreter must be used to translate and become a mediator of the patient and listener, which adds another degree of separation from the patient.

Tracheostomy and Communication

A tracheostomy is a surgical procedure that patients get when there is an issue with their airway which can result from an acquired injury or a developmental disability. The Mayo Clinic

(2023) defines a tracheostomy as an incision made at the front of the neck where a hole (stoma) is created into the windpipe, otherwise known as the trachea. A tracheostomy tube is placed into the open hole to allow breathing or to connect a ventilator to the patient if they are unable to breathe independently. Trach tubes may be cuffed or cuffless, referring to a balloon like feature at the distal end of the tube (Tracheostomy Education, 2019). With cuffed trach tubes, there are typically three types: high volume-low pressure cuffs, low volume-high pressure cuffs, and foam filled cuffs. All of the cuff types serve the same purpose, with the intention of blocking off and preventing airflow to the upper respiratory tract and out of the tube. The cuff can be deflated, allowing the airflow to pass through the upper airway system. This is essentially how we speak, by allowing air from the lungs to pass through the trachea, the larynx, and out of the mouth. Both cuffed and cuffless tubes may also be fenestrated or non-fenestrated. A fenestrated tube simply means there is a cut/opening on the tube, allowing for an increased amount of air to pass through the trachea.

Due to the invasive nature and location of a tracheostomy, it can be a potentially life-altering procedure where patients often have trouble with a significant basic human function such as communication (Tolotti et al., 2018; Grossbach et al., 2011; McGrath et al., 2016; Pandian et al., 2022). A study done on patients who received a tracheostomy identified themes and factors relating to their experience with communication in the ICU (Tolotti et al., 2018). The authors discussed the patient's feelings of powerlessness and frustration as well as experiencing misunderstanding, resignation, and anger. These patients also experienced discomfort with not knowing what was happening in the context of their medical stay, feeling like other people gave up on them, living in isolation, and feeling invisible (Tolotti et al., 2018).

Speaking with a Tracheostomy

Speaking with a tracheostomy is generally unlikely because air that is inhaled, either naturally or mechanically using a ventilator, does not pass through the vocal cords. While less likely, it is not impossible to speak with a tracheostomy. When a trach tube cuff is deflated, air is able to pass around the tube, through the larynx, and out of the mouth, producing speech known as "leak speech" (Grossbach et al., 2011, p. 55; McGrath et al., 2016, p. 23). However, the amount of air that passes through the upper airway when leak speech occurs is typically far less than the amount of air necessary to speak at a natural level and may need to increase the level of airflow (McGrath et al., 2016, p. 23).

Grossbach et al. (2011) gave an overview of different modifications to allow for patients to speak using a one-way speaking valve or types of tracheostomy tubes using ventilator air and natural air. A speaking valve is typically used with tracheostomy patients to produce more airflow through the upper airway and allowing for more natural speech. McGrath et al. (2016) discussed the potential of adding a fenestration, or a cut to a part of the trach tube enabling an increased airflow to pass through the airway.

Once a patient weans off the assistive methods of speech production, by using a speaking valve or cuff deflation, decannulation (the removal of a tracheostomy tube) is the ideal and closest result to natural speech production.

Above Cuff Vocalization

Systematic and literature reviews on above cuff vocalization (McGrath et al., 2016; Petosic et al., 2021; Mills et al., 2022b) have examined the safety and feasibility (McGrath et al., 2019) and discussed the prevalence, implementation approaches, and opinions (Mills et al., 2022a) of this alternative method of speech production.

One study explored a novel technique for communication with tracheostomy and ventilator dependent patients using Above Cuff Vocalization (McGrath et al., 2016). Typically, trach patients have cuffed tubes that inflate and deflate. Some trach tubes have a subglottic suction tube that is used to remove secretions that gathers above the cuff (Tracheostomy Education, 2020). McGrath et al. discussed the method of "directing a retrograde flow of gas via the Subglottic Suction (SGS) tube to exit above the cuff" and therefore summarizing that "patients can theoretically vocalize" (p. 20). In other words, the subglottic suction tube would potentially serve a dual purpose of removing secretions and directing airflow above the cuff, resulting in vocalization.

Some studies outlined potential limitations of ACV. Petosic et al. (2021) complications included burping, subglottic air trapping, retching, nausea, cuff breakage, tracheal distention, and subcutaneous emphysema (p. 23). Other outcomes such as patient discomfort, excessive secretions, excessive coughing and stomal air leaks were some of the most common complications reported (McGrath et al., 2019; Petosic et al., 2021; Mills et al., 2022a).

Augmentative Technology

Haring et al. (2022), a medical team from the Department of Otolaryngology out of Michigan, studied the effects of using augmentative technology on communicating and the quality of life with patients who received a tracheostomy. With the modernization and integration of smart devices, patients have a wide range of access to technology. Text-to-speech functions, speech generating devices/applications, and communication boards using pictures and symbols are some examples of augmentative and alternative communication (AAC). Furthermore, Haring et al. (2022) discussed the significance of AAC technology on the quality of

life of tracheostomy patients, suggesting the positive impact that alternative communication methods may have on patients who struggle to verbally communicate.

Identity

Patients who at one point in their life were able to speak naturally but have an emergency or necessary tracheostomy procedure may find immense frustration, anxiety, and stress from losing their voice (Tolotti et al., 2018; Grossbach et al., 2011; Haring et al., 2022). In addition to the discomfort and loss of ability to speak using a natural voice that comes with an invasive surgery like a tracheostomy, many patients experience a wide range of additional hardships that take away from a very important piece of their identity: their voice.

Comorbidity with Speech and Language Disorders

In addition to the surgery itself, some tracheostomy patients may carry comorbid diagnoses of speech/language disorders. Speech and language disorders can be developmental or acquired from injury/illness. According to the American Speech-Language-Hearing Association (2023), speech and language disorders such as apraxia, dysarthria, stuttering, aphasia as well as medical conditions such as dementia, cancers, or brain injuries may cause issues in verbal communication. Some of the resulting problems may be physical/motor impairments that affect speech, whereas some may be neurological/psychological that affect language.

ASHA (2023) describes apraxia of speech (AOS) as a "motor speech neurological condition where there is difficulty with speech from a coordination standpoint. The brain has difficulty moving the physical components to create the necessary sounds, but they are cognitively aware of what they want to say". Dysarthria is a speech disorder that is caused by injury, damage, and/or paralysis often resulting in muscle weakness. Stuttering occurs when a

patient knows what they want to say but the resulting speech is a combination of cluttered, disrupted, repeated, or prolonged sounds, syllables, or words. Aphasia is a condition that is acquired when areas of the brain are damaged, leading to impaired language understanding or expression. Broca's area of the brain, located in the inferior frontal gyrus near the motor cortex is responsible for speech production, planning and organizing speech movements, and expressive language. When Broca's area is impaired, called Broca's aphasia, a patient is able to understand and comprehend speech but has difficulty with finding the words to speak. Wernicke's area is located in the posterior superior temporal gyrus and when impaired, can lead to difficulty understanding words and meanings.

For patients who have a speech and language disorder in addition to their tracheostomy procedure, verbal communication is difficult, and they require other methods of communication.

Anatomy of the Voice

The human voice is an intricate system of parts that work together to produce sounds and expressions. In *Music and Sound in the Healing Arts* (1987), John Beaulieu highlights the voice and its parts. He categorizes the anatomy of the voice into the power source, the vibrators, the articulators, and the resonators, with the brain functioning as the head of operations, controlling the speech processes, production, and perception.

The brain is comprised of the left and right hemispheres, both of which are responsible for different functions. As outlined in the speech disorders section, areas such as Broca's area and Wernicke's area have a direct impact on speech production and understanding. Our brain also has areas that are indirectly related to speech such as memory, emotions, and fear.

The power source is what is used to supply the breath stream by taking in and pushing out air from our body. The diaphragm, lungs, trachea, and bronchi make up the power source. Air from the lungs is pushed upwards and outwards by the support of the diaphragm and through the trachea and bronchi. The larynx and vocal cords shape and form the airflow which are used to modulate the breath stream. Air travels to the upper airway where the lips, teeth, tongue, gum ridge, hard palate, soft palate, and uvula fine-tune the breath into a more specific and distinct output. Finally, the nasal cavities, oral cavity, throat cavity (pharynx), and the chest cavity, known as the resonators, are used to select and suppress overtones and amplify or restrain the final produced sound.

Ten Ingredients of the Voice

Loudness, pitch/note, pitch fluctuation, register, harmonic timbre/resonance, nasality (violin), free air, attack, disruption, and articulation are listed as "ten ingredients of the voice" adapted from Paul Newham's *Using Voice and Song in Therapy* (1999). Loudness is perceived on a spectrum from loud to quiet, measured by decibels. The pitch or note is perceived on a spectrum from low frequencies to high frequencies, measured in hertz. Pitch fluctuation is perceived as fast or slow or great or small, referring to the change of notes. The register of the voice refers to the range of tones that is described as modal, falsetto, or blended. Other common phrases to describe the vocal register are chest voice, head voice, and mixed voice. Some also include vocal fry and whistle registers. Harmonic timbre describes the tonal quality of sound. For example, a flute is described as having a high timbre, a clarinet has middle timbre, and a saxophone has a low timbre. The nasality and free air relate to how air interacts with our voice and is described from having minimum to maximum amounts. The attack describes the initial harshness of the voice that is measured from lesser to greater. The disruption refers to the

amount of friction involved in the vocal output, from mild to severe. Finally, articulation is the detailed shaping and forming of the voice using consonant and vowel shapes.

Overlap of Music and Speech/Communication

Music is often described as a universal language, referring to the widespread cultural phenomenon that can be found around the world throughout history. As such, there are many similarities and overlaps between music and speech including, but not limited to: pitch, prosody, rhythm, melody, intonation, amplitude, syntax, violations, and mechanics.

While music from different cultures may vary such as the use of multi-tonal scales, there are some factors that are shared. Music, at its core, can be described as an arrangement of sounds that create an expression of melody, harmony, and rhythm. This oversimplified notion allows two people to engage in making music together even though they don't communicate using the same spoken language.

Although there are localized areas of the brain that control our speech and/or language, there is no singular, specific region of our brain that is reserved for music. When people engage with musical qualities such as rhythm or melody, there is substantial neural overlap between music and speech processing (Peretz et al., 2015). The areas of the brain that process speech and the areas that process music overlap with each other due to shared qualities that can be recognized in both speech and music such as prosody, intonation, or amplitude. An example of the significance of inflection and intonation of speech is how we can distinguish tonal languages from nontonal languages even if we are not fluent in one or the other. Patel et al. (2006) share "the prosody of a culture's native language is closely associated with the rhythms and melodies

of its instrumental music", introducing an interesting overlap of a culture's language and their music.

Another shared component of both speech and music is the mechanics of sound produced in a physical sense. In its essence, speech and music can be described as sound waves. When a human sings or speaks, the required mechanisms for both actions include the inhalation of air into the lungs, pushing out using the diaphragm, vibrating the vocal cords, and forming consonants/vowels with the mouth and tongues.

Music Therapy and Speech Pathology

Music therapy is the clinical & evidence-based use of music interventions to accomplish individualized goals within a therapeutic relationship by a credentialed professional who has completed an approved music therapy program (AMTA, 2023). In music therapy, the voice is a vital instrument used for: re-creating songs, improvisation, composing songs for the voice, and receptive music (Knight et al., 2018). For patients who are unable to use their voice to the fullest due to impairment/damage or developmental delay, music therapy can be used as a communicative medium. For instance, patients with aphasia use Neurologic Music Therapy (NMT) techniques including Melodic Intonation Therapy (MIT), Musical Speech Stimulation (MUSTIM), and Rhythmic Speech Cueing (RSC) (Thaut and Hoemberg, 2014) which will be detailed in the next section. Of the many NMT techniques, Oral Motor and Respiratory Exercises (OMREX) would be an appropriate intervention for tracheostomy patients.

Speech Language Pathologists (SLP) work to prevent, assess, diagnose, and treat speech, language, social communication, cognitive-communication, and swallowing disorders in children and adults (ASHA). Speech Pathologists also utilize Melodic Intonation Therapy among other

techniques. SLP's generally use a lot of repetition in most of their techniques. For patients with apraxia, speech sounds and phonemic segmentation is the focus. Dysarthria patients use breathing exercises (using an incentive spirometer), prosody/intonation exercises, and a slower speech rate. Stuttering patients use easy onset techniques as well as light contact techniques for harsh stutters.

Current Music Therapy Techniques for Communication and Speech Disorders

Much of the music therapy literature related to communication and speech disorders discusses the effectiveness of nonverbal communication using musical interventions. Music therapy techniques such as fill-in-the-blank exercises for goals in socialization are utilized for people with autism spectrum disorder (Geist et al., 2008).

Melodic Intonation Therapy (MIT) is an NMT technique that is used for expressive aphasia patients "to facilitate spontaneous and voluntary speech through sung and chanted melodies which resemble natural speech intonation patterns" (Thaut, 2005). It emphasizes the usage of "rhythmic elements to engage language-capable regions of the undamaged right hemisphere" (Thaut et al., 2014, p. 140).

Musical Speech Stimulation (MUSTIM) utilizes "musical materials such as songs, rhymes, chants, and musical phrases to simulate prosodic speech gestures and trigger automatic speech" (Thaut, 2014, p. 146). MUSTIM is typically used with patients who experienced a left hemisphere stroke or brain injury.

Mainka and Mallien (2014) define rhythmic speech cueing (RSC) as using speech rate control via auditory rhythm to "improve temporal characteristics such as fluency, articulatory rate, pause time, and intelligibility of speaking. Tempo is the most important factor for the

therapeutic power of the technique. There are two modes of cueing speech production: metric and patterned. Metric cueing is a pulsed auditory stimulation where the patient is asked to match either one syllable or one full word to one beat. In patterned cueing, the patient reproduces a pre-structured rhythmic sentence at a given tempo (p. 150).

Oral Motor and Respiratory Exercises (OMREX) is an NMT technique that uses musical materials and exercises to address the improvement of "articulatory control, respiratory strength, and function of the speech apparatus" (Mertel, 2014, p. 161). Mertel (2014) also outlines wind instruments such as the recorder, the harmonica, the melodica, and the kazoo that are used for breathing exercises in order to "improve vocal strength, exercise laryngeal function, increase respiratory capacity, and refine oral motor function (p. 170). Other areas that these instruments are used for include: "maintaining mouth closure, coordination of breathing patterns, supporting diaphragmatic breathing, and strengthening/promoting the use of the voice" (p. 170).

Current Medical Interventions for Communication and Speech Disorders

Electrolarynx

Some patients who have damage to the vocal cords or voice box (larynx) may use an electrolarynx as a tool to speak. An electrolarynx is a battery powered device that operates by producing a constant pitch that can be adjusted for male and female speakers. The electrolarynx is most commonly placed at the neck of a patient and a button is pressed to produce a vibrated electronic sound that is transmitted through the neck and creates speech using articulators such as "the lips, teeth, tongue, jaw, and velum" (Liu and Ng, 2007).

A type of intra-oral electrolarynx modification may be used to transmit the sound directly into the patient's mouth, resulting in less energy leakage and better speech quality compared to

the neck type (Liu and Ng, 2007). One of the biggest challenges with electrolarynxes is controlling the pitch. A finger-controlled type is commercially available, where the patient is able to control the pitch of the device using a pressure-sensitive button. Other types of pitch-changing types are the expiration-control type and electromyographic (EMG)-control type.

In Liu and Ng's discussion of electrolarynxes, two aspects of improving the speech quality are mentioned. First is the ability to adjust pitch and intensity in real-time during phonation. Second is the integration of electronic media for communication in an ever-advancing digital world.

Talk Box

The talk box is an instrument effect device that was developed in 1939 and has primarily been used in musical performance. In its simplest form, the configuration of a talk box consists of a mini speaker amplifier with an input for instruments and an output where a plastic tube can be attached. The distal end of the plastic tube is placed in a speaker's mouth where they can use their articulators to form consonant and vowel shapes. The resulting effect amplifies the instrument's output into the speaker's mouth where words or phrases can be spoken or sung.

Typically, instruments such as the guitar or keyboard are used with a talk box. It is encouraged to play single notes rather than chords to achieve the best resulting sound, with minimal distortion or clashing of notes.

Anatomy of Voice and Interaction with a Talk Box

When speaking, air is inhaled into the lungs, pushed to the upper airway with support from the diaphragm, passing the larynx, and finally out of the mouth where the lips, teeth, and tongue shape the final output into specific expressions. For tracheostomy patients, mechanically

ventilated air or naturally inhaled air is not able to pass through the upper airway and through the vocal cords. So even though a tracheostomy patient may have functioning and intact articulators or resonators, the power source and vibrators are primarily lacking.

The talk box is able to function as both the power source and the vibrators. Sound that is played from an instrument is amplified and pushed through the tube, acting as an elongated speaker. As sound is simply summarized as waves that vibrate at a frequency, the sound that is produced using a talk box can essentially take the place of the vibrating air that is produced when air is pushed through the vocal cords.

Comparison of Talk Box with Current Electrolarynx

The commonly used electrolarynx has similarities to a talk box where the function of both devices is to produce an external sound that is modified by the patient. The electrolarynx is battery powered and portable, making it convenient to bring around with the patient. It operates by placing the device on the neck and pressing a button to produce a monotonous, robotic tone. The frequency of the tone can be adjusted for male and female speakers. There is an intra-oral modification that can be used with an electrolarynx, where instead of placing the device directly on the patient's neck, a small, straw-like tube is placed in the patient's mouth and allowing sound to be directed to the mouth's resonators and articulators to form words and phrases.

The talk box functions closely to an intra-oral electrolarynx. A larger tube is placed in the patient's mouth as it transfers an amplified sound from an instrument being played. However, the talk box offers more flexibility and a wider range of tones compared to an electrolarynx. The talk box also addresses the two aspects of improving speech quality of using an electrolarynx that Liu and Ng (2007) described.

First, the ability to adjust pitch and intensity in real-time during the patient's phonation is addressed in using a keyboard with the talk box. A keyboard has upwards of 88 keys to be played as well as a plethora of digital instruments to choose from. There are many sounds that can be played using a keyboard compared to the singular tone electrolarynx. In addition to the pitch and intensity/loudness variables being addressed, the use of a musical instrument such as a keyboard or guitar allows other ingredients of the voice including pitch fluctuation, register, harmonic timbre/resonance, nasality, attack, and disruption to be factored into the patient's phonation.

Second, with the constantly developing musical world, the aspect of integrating electronic media for communication can be addressed with the talk box and a MIDI keyboard/controller. MIDI, or musical instrument digital interface, allows a user to map different sounds, samples, beats, and chords to a single key on a keyboard or a pad on a launchpad instrument. This allows a seemingly infinite number of options to map to different keys as it can be updated and modified instantly. If a key on a MIDI keyboard is mapped to a specific commonly used phrase such as "It's nice to meet you", the patient would be able to mouth the words and have the instrument "speak" for them, while creating an experience for the patient to feel as if they are speaking with their own voice. A modified version of this example could map the pitch, inflection, tempo, and other musical qualities to a key without an actual recording of the phrase "It's nice to meet you", allowing the patient to use their own articulators to create a more natural speaking experience.

The electrolarynx is the size of a small, portable flashlight which is very convenient for the patient. The cost of an electrolarynx ranges from $340 to $2500 which is a very expensive

purchase for a patient and their family. It also costs additional amounts for intra-oral modifications and a wide variety of companies to choose from.

The talk box is the size of a guitar effects pedal, powered by an 18v power supply, and is accessible to purchase online or at a music store. The cost of a talk box unit ranges from $120 to $230 to purchase. This method of integrating a talk box in music therapy is also non-invasive and non-permanent.

Discussion

Tracheostomy patients struggle with one of the most significant and essential components to the human identity, that is, communication. Many patients who get a tracheostomy procedure after the age for fully developed verbal language skills experience difficulties such as increased frustration, stress, and anxiety (Tolotti et al., 2018; Grossbach et al., 2011; Haring et al., 2022). Speech assistance for tracheostomy patients currently include above cuff vocalization efforts using one-way speaking valves (Grossbach et al., 2011), augmentative and alternative communication methods such as speech generating devices and the electrolarynx (Liu and Ng, 2007).

As mentioned in Liu and Ng's (2007) discussion of electrolarynxes, two aspects of improving speech quality are mentioned. First is the ability to adjust pitch and intensity in real-time during phonation. Second is the integration of electronic media for communication in an ever-advancing digital world. The literature discusses the current electrolarynx offering a simple, minimal method of achieving speech by placing the device on the patient's neck to produce a monotonous, robotic tone when pressed. While this device and method offers a quick fix solution by being favorable in accessibility and simplicity, it lacks the more natural sounding tonal and dynamic qualities of speech. Another important component to the electrolarynx is the

affordability and capitalistic factors, where there are countless brands to select from and having an average cost that ranges from several hundreds to thousands of dollars.

Another important consideration is the identity factor that is lost in tracheostomy patients. Patients who at one point in their life had no issues with speaking who undergo an emergency or unexpected tracheostomy surgery may have to grieve the loss of their unique voice. What was once something directly integrated into an individual's identity and something that they carried with them everywhere they went can be eliminated due to the invasive and life-altering procedure. An individual's unique identity is not something that can be simply replaced with a single-toned electronic sound.

Comparisons can be made between the intraoral electrolarynx and the talk box which function in a very similar manner. Both methods require the patient to place a tube or straw-like piece in the mouth, shape the speech output with their mouth, and allow the device to produce the sound. Some significant differences in the intraoral electrolarynx and the talk box is the variability and control that a talk box offers.

Application in Music Therapy

As previously mentioned, using an instrument with the talk box allows for a real-time adjustment of pitch, quality, loudness, timbre, and tempo. A large part of music therapy incorporates the use of music and instruments to help a patient with non-musical goals. If a patient is able to play an instrument on their own, it would encourage autonomy and an individual identity to the speaking/singing experience. If a patient is unable to play an instrument proficiently, the music therapist, having competency in guitar and piano, would assist the speaking/singing experience.

Due to the significant overlap in speech and music, the use of an instrument to "play" the notes of verbal speech offers a more natural sounding experience compared to the electrolarynx. In addition to a speaking experience, the patient also has an opportunity to sing. While figuring out the specific notes to play to mimic a talking phrase may be overwhelming, a singing experience is much simpler. A song's melody is generally much simpler to play on an instrument compared to the variability of pitch, tempo, and loudness when talking. A song's melody is also available to read on a musical chart or accessible from the patient's memory. If a patient has a favorite song they enjoy listening to, they can also "sing" along by playing the notes of the melody on the instrument and using the talk box.

Conclusion

This thesis reviews the present literature that is relevant to communication, speech devices, identity, and treatment methods for tracheostomy patients, including a novel music therapy theoretical approach. Research suggests the exploration of alternative methods to the currently used electrolarynx and one-way speaking valve to provide more natural sounding speech.

There is an underwhelming amount of medical research or clinical implications for using a talk box as a method for speech therapy or music therapy. Peer-reviewed journal articles about electrolaryngeal speech are limited and is encouraged to further research.

Integrating a musical approach to speech is commonly used in neurologic music therapy techniques such as Melodic Intonation Therapy (MIT), Musical Speech Stimulation (MUSTIM), Rhythmic Speech Cueing (RSC), and Oral Motor and Respiratory Exercises (OMREX) (Thaut

and Hoemberg, 2014). Research on music therapy and its impact on decreasing stress and anxiety is very accessible and suggests positive results.

As the name implies, the talk box could provide tracheostomy patients with a non-invasive and non-permanent approach to speaking and singing, as well as an opportunity to reclaim a missing part of their identity that cannot be replaced or replicated.

References

American Music Therapy Association (2023) *American Music Therapy Association (AMTA)*. https://www.musictherapy.org/

American Speech-Language-Hearing Association. (2023). *Speech and Language Disorders*. https://www.asha.org/public/speech/disorders/

Blischak, D., Lombardino, L., & Dyson, A. (2003). Use of Speech-Generating Devices: In Support of Natural Speech. *Augmentative and Alternative Communication*, *19*(1), 29–35. https://doi.org/10.1080/0743461032000056478

Federman, M. (2004). *What is the Meaning of the Medium is the Message?* http://individual.utoronto.ca/markfederman/MeaningTheMediumistheMessage.pdf

Geist, K., McCarthy, J., Rodgers-Smith, A., & Porter, J. (2008). Integrating Music Therapy Services and Speech-Language Therapy Services for Children with SevereCommunication Impairments: A Co-Treatment Model. *International Journal of Convergence in Healthcare*, *1*(1). https://doi.org/10.55487/ijcih.v1i1.6

Grossbach, I., Stranberg, S., & Chlan, L. (2011). Promoting Effective Communication for Patients Receiving Mechanical Ventilation. *Critical Care Nurse*, *31*(3), 46–60. https://doi.org/10.4037/ccn2010728

Haire, N., & MacDonald, R. (2021). Understanding how humour enables contact in music therapy relationships with persons living with dementia: A phenomenological arts-based reflexive study. *The Arts in Psychotherapy*, *74*, 101784. https://doi.org/10.1016/j.aip.2021.101784

Haring, C. T., Farlow, J. L., Leginza, M., Vance, K., Blakely, A., Lyden, T., Hoesli, R. C., Neal, M. E. H., Brenner, M. J., Hogikyan, N. D., Morrison, R. J., & Casper, K. A. (2022). Effect of Augmentative Technology on Communication and Quality of Life After Tracheostomy or Total Laryngectomy. *Otolaryngology–Head and Neck Surgery*, *167*(6), 985–990. https://doi.org/10.1177/01945998211013778

Holler, J., & Levinson, S. C. (2019). Multimodal Language Processing in Human Communication. *Trends in Cognitive Sciences*, *23*(8), 639–652. https://doi.org/10.1016/j.tics.2019.05.006

Knight, A., LaGasse, B., & Clair, A. (2018). *Music Therapy: An Introduction to the Profession*.

Krøier, J. K., Stige, B., & Ridder, H. M. (2021). Non-Verbal Interactions Between Music Therapists and Persons with Dementia. A Qualitative Phenomenological and Arts-Based Inquiry. *Music Therapy Perspectives*, *39*(2), 162–171. https://doi.org/10.1093/mtp/miab008

Lapum, J., St-Amant, O., Hughes, M., & Garmaise-Yee, J. (2020). Communcation is Learned. In *Introduction to Communication in Nursing*. https://pressbooks.library.torontomu.ca/communicationnursing/chapter/communication-is-learned/

Liu, H., & Ng, M. L. (2007). Electrolarynx in voice rehabilitation. *Auris Nasus Larynx*, *34*(3), 327–332. https://doi.org/10.1016/j.anl.2006.11.010

Luhmann, N. (1992). What is Communication? *Communication Theory*, *2*(3), 251–259. https://doi.org/10.1111/j.1468-2885.1992.tb00042.x

Mainka, S., & Mallien, G. (2014). Rhythmic Speech Cueing (RSC). In M. H. Thaut & V. Hoemberg, *Handbook of Neurologic Music Therapy* (pp. 150–160).

McGrath, B. A., Wallace, S., Wilson, M., Nicholson, L., Felton, T., Bowyer, C., & Bentley, A. M. (2019). Safety and feasibility of above cuff vocalisation for ventilator-dependant patients with tracheostomies. *Journal of the Intensive Care Society*, *20*(1), 59–65. https://doi.org/10.1177/1751143718767055

McGrath, B., Lynch, J., Wilson, M., Nicholson, L., & Wallace, S. (2016). Above cuff vocalisation: A novel technique for communication in the ventilator-dependent tracheostomy patient. *Journal of the Intensive Care Society*, *17*(1), 19–26. https://doi.org/10.1177/1751143715607549

McLuhan, M. (1964). *Understanding Media*.

Mertel, K. (2014). Oral Motor and Respiratory Exercies (OMREX). In M. H. Thaut & V. Hoemberg, *Handbook of Neurologic Music Therapy* (pp. 161–178).

Mills, C. S., Michou, E., Bellamy, M. C., Siddle, H. J., Brennan, C. A., & Bojke, C. (2022). Determining the Prevalence, Implementation Approaches, and Opinions of Above Cuff Vocalization: A Survey of Health Care Professionals. *Archives of Physical Medicine and Rehabilitation*, *103*(3), 394–401. https://doi.org/10.1016/j.apmr.2021.08.016

Mills, C. S., Michou, E., King, N., Bellamy, M. C., Siddle, H. J., Brennan, C. A., & Bojke, C. (2022). Evidence for Above Cuff Vocalization in Patients With a Tracheostomy: A Systematic Review. *The Laryngoscope*, *132*(3), 600–611. https://doi.org/10.1002/lary.29591

O'Donoghue, J., Moss, H., Clements-Cortes, A., & Freeley, C. (2020). Therapist and individual experiences and perceptions of music therapy for adolescents who stutter: A qualitative exploration. *Nordic Journal of Music Therapy*, *29*(4), 353–370. https://doi.org/10.1080/08098131.2020.1745872

Pandian, V., Hopkins, B. S., Yang, C. J., Ward, E., Sperry, E. D., Khalil, O., Gregson, P., Bonakdar, L., Messer, J., Messer, S., Chessels, G., Bosworth, B., Randall, D. M., Freeman-Sanderson, A., McGrath, B. A., & Brenner, M. J. (2022). Amplifying patient voices amid pandemic: Perspectives on tracheostomy care, communication, and connection. *American Journal of Otolaryngology*, *43*(5), 103525. https://doi.org/10.1016/j.amjoto.2022.103525

Patel, A. D., Iversen, J. R., & Rosenberg, J. C. (2006). Comparing the rhythm and melody of speech and music: The case of British English and French. *The Journal of the Acoustical Society of America*, *119*(5), 3034–3047. https://doi.org/10.1121/1.2179657

Peretz, I., Vuvan, D., Lagrois, M.-É., & Armony, J. L. (2015). Neural overlap in processing music and speech. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1664), 20140090. https://doi.org/10.1098/rstb.2014.0090

Petosic, A., Viravong, M. F., Martin, A. M., Nilsen, C. B., Olafsen, K., & Berntzen, H. (2021). Above cuff vocalisation (ACV): A scoping review. *Acta Anaesthesiologica Scandinavica*, *65*(1), 15–25. https://doi.org/10.1111/aas.13706

Rispoli, M. J., Franco, J. H., van der Meer, L., Lang, R., & Camargo, S. P. H. (2010). The use of speech generating devices in communication interventions for individuals with developmental disabilities: A review of the literature. *Developmental Neurorehabilitation*, *13*(4), 276–293. https://doi.org/10.3109/17518421003636794

Silverman, M. J., & Bibb, J. (2018). Acute care mental health workers' assumptions and expectations of music therapy: A qualitative investigation. *The Arts in Psychotherapy, 59*, 94–100. https://doi.org/10.1016/j.aip.2018.05.002

Thaut, C. P. (2014). Musical Speech Stimulation (MUSTIM). In M. H. Thaut & V. Hoemberg, *Handbook of Neurologic Music Therapy* (pp. 146–149).

Thaut, M. H. (2005). Neurologic Music Therapy Techniques and Definitions. In *Rhythm, Music and the Brain*. New York and London: Taylor and Francis Group.

Thaut, M. H., & Hoemberg, V. (2014). *Handbook of Neurologic Music Therapy*. Oxford University Press.

Thaut, M. H., Thaut, C. P., & McIntosh, K. (2014). Melodic Intonation Therapy (MIT). In M. H. Thaut & V. Hoemberg, *Handbook of Neurologic Music Therapy* (pp. 140–145).

Tolotti, A., Bagnasco, A., Catania, G., Aleo, G., Pagnucci, N., Cadorin, L., Zanini, M., Rocco, G., Stievano, A., Carnevale, F. A., & Sasso, L. (2018). The communication experience of tracheostomy patients with nurses in the intensive care unit: A phenomenological study. *Intensive and Critical Care Nursing, 46*, 24–31. https://doi.org/10.1016/j.iccn.2018.01.001

Tracheostomy Education. (2019, November 2). *Cuffed versus Cuffless Tracheostomy Tubes*. Tracheostomy Education. https://tracheostomyeducation.com/cuffed-versus-cuffless-tracheostomy-tubes/

van der Meer, L. A. J., & Rispoli, M. (2010). Communication interventions involving speech-generating devices for children with autism: A review of the literature. *Developmental Neurorehabilitation, 13*(4), 294–306. https://doi.org/10.3109/17518421003671494

THESIS APPROVAL FORM

Lesley University

Graduate School of Arts & Social Sciences Expressive Therapies Division

Master of Arts in Clinical Mental Health Counseling: Music Therapy, MA

Student's Name: Jonathan Kang

Type of Project: Thesis

Title: Talk Box in Music Therapy with Speech and Language Impairments Resulting From Tracheostomy: A Critical Review of the Literature

Date of Graduation: May 2024

In the judgment of the following signatory this thesis meets the academic standards that have been established for the above degree.

Thesis Advisor: Dr. Rebecca Zarate MT-BC, AVPT, LCAT