The BCH message banking process™, voice banking, and double-dipping™

John Costello & Martine Smith

◌ View supplementary material ◲

▦ Published online: 08 Jan 2022.

◰ Submit your article to this journal ◲

▥ Article views: 117

◷ View related articles ◲

◉ View Crossmark data ◲

Taylor & Francis
Taylor & Francis Group

Check for updates

RESEARCH ARTICLE

The BCH message banking process™, voice banking, and double-dipping™

John Costello[a]* 🆔 and Martine Smith[b] 🆔

[a]Augmentative Communication Program and Jay S. Fishman ALS Augmentative Communication Program, Boston Children's Hospital, Adjunct Faculty Boston University, Boston, MA, USA; [b]Department of Clinical Speech and Language Studies, Trinity College Dublin, Dublin, Ireland

ABSTRACT

Significant advances have been made in interventions to maintain communication and personhood for individuals with neurodegenerative conditions. One innovation is Message Banking, a clinical approach first developed at Boston Children's Hospital (BCH). This paper outlines the Message Banking process as implemented at BCH, which includes the option of "Double Dipping," where banked messages are mined to develop personalized synthesized voices. More than a decade of experience has led to the evolution of six core principles underpinning the BCH process, resulting in a structured introduction of the associated concepts and practices with people with amyotrophic lateral sclerosis (ALS) and their families. These principles highlight the importance of assigning ownership and control of the process to individuals with ALS and their families, ensuring that as a tool it is empowering and offers hope. Changes have been driven by feedback from individuals who have participated in the BCH process over many years. The success of the process has recently been extended through partnerships that allow the recorded messages to be used to develop individual personalized synthetic voices to complement banked messages. While the process of banking messages is technically relatively simple, the full value of the process should be underpinned by the values and principles outlined in this tutorial.

Over the past decade, many strategies have evolved to support individuals with neurodegenerative conditions to maintain their communication status and persona (Fried-Oken, Mooney, & Peters, 2015). These innovations include hardware developments in speech-generating devices (SGDs), enhancements in the quality and range of synthetic speech, facilities to support personalization of synthetic speech through voice banking (e.g. Bunnell & Pennington, 2010), and solutions to support individualized banking of personally meaningful messages, through message banking (Costello, 2000). Since it was originally described, the concept of banking personal messages and individual voices has been applied in different ways. This paper (a) contrasts message banking with voice banking; and (b) describes the philosophy and practice of the BCH Message Banking process, an innovation that combines features of both voice banking and message banking to support the creation of personalized synthetic voices and includes "double dipping," now a routine option when message banking within the BCH model.

The role of personal voice

For people with a neurodegenerative disease such as ALS, speech loss is difficult because of not only the dramatic

change to communication independence but also the threatened loss of identity (Cave & Bloch, 2021). Just as each individual's fingerprint reveals their unique biological identity, voice is an acoustical fingerprint (Costello, 2000), central "in the expression and continual (re)construction of personal and social identity, and in forming and maintaining social bonds" (Nathanson, 2017, p. 75). Voice also carries core information about emotional state and communicative intent (Juslin, Laukka, & Bänziger, 2018). "Talkers speak to be understood, and understanding means more than intelligibility" (Pardo, 2012, p. 764). Inferring intent involves considering *who* is speaking, *what* is being said and *how* it is being expressed, the multi-layered information expressed through tone, volume, intonation, and stress. Voice reveals emotion (humor, despair, sadness) as well as a speaker's personal state (e.g. nervousness) and health (Podesva & Callier, 2015), regardless of the words spoken (e.g. Juslin et al., 2018).

The unique role of familiar voices

The role of voice as a marker of identity is important for not only the speaker but also listeners. Although voices are all around us, familiar voices are special (e.g. Sidtis & Kreiman, 2012) in their evolutionary value, diffuse neurological

representation (Maguinness, Roswandowitz, & von Kriegstein, 2018; McGettigan & & others, 2015), and in the unique impact they have on listeners who have a personal relationship to the speaker (Seltzer, Ziegler, & Pollak, 2010). Sidtis and Kreiman (2012, p. 154) describe the personal, familiar voice as an "emotional auditory event." Hearing a mother's voice, even by phone, can be similar in its effect in reducing a daughter's stress levels (as measured by levels of oxytocin) to having direct physical contact (Seltzer et al., 2010), unlike the impact of instant message contact between mothers and daughters (Seltzer, Prososki, Ziegler, & Pollak, 2012). These findings suggest that the unique features of an individual's voice help to nurture and sustain social bonds and relationships. Loss of voice, therefore, impacts not only the individual speaker but also has the potential to deprive others of emotional connectedness.

Voice in the context of neurodegenerative disease processes

Many disease processes (e.g. Huntington's disease, Parkinson's disease, multiple sclerosis) can impact speech intelligibility. In ALS, as many as 95% of people become unable to communicate using natural speech at some point in the disease progression (Beukelman, Fager, & Nordness, 2011; Pattee et al., 2019). The time from onset of initial symptoms to significantly reduced intelligibility varies greatly across individuals and across types of ALS (Ball, Beukelman, & Pattee, 2002). A Finnish study of 30 participants with bulbar onset ALS (Makkonen, Ruottinen, Puhto, Helminen, & Palmio, 2018) found that speech remained adequate on average for 18 months following the first bulbar symptoms, but time ranged from 6 to 40 months. Given the high likelihood in ALS of loss of speech over a relatively short period of time, it is vitally important to identify strategies as early as possible to preserve existing abilities and identify potential options for augmentative and alternative communication (AAC) solutions before intelligibility becomes a critical concern. Two technology-supported strategies that address the latter aim are voice banking and message banking.

Preserving voice through voice banking and message banking

Voice banking involves using high-quality digital recordings of an individual's speech and segmenting those recordings to generate a unique phonetic database. These phonetic segments can then be combined in novel sequences to create synthetic utterances with personalized synthesized voice that can be accessed through text-to-speech (Bunnell & Pennington, 2010; Pullin, Treviranus, Patel, & Higginbotham, 2017). To complete this process, individuals register with a voice banking service and, using a computer with internet connection, record in a quiet environment with a high-quality microphone. Depending on the service, 1600 or more phrases may be needed to establish a sufficiently large data set to support the creation of a personal voice. In recent years, many voice banking services have applied deep neural

network (DNN) technology, a form of machine learning where technologies exploit aspects of artificial intelligence to classify and order information in ways that go beyond simple input/output protocols (Hinton et al., 2012). This allows for a reasonable quality personalized voice to be created from samples as small as 50 recordings. Despite these advances, generating the necessary quantity of recordings can be challenging for individuals already experiencing speech deterioration. Voice banking is an important development in recognizing the unique and specific role of personal voice in maintaining identity; however, apart from the physical demands of recording, even high-quality voices lack the facility to generate the prosody that characterizes natural speech.

Message banking (Costello, 2000) is a clinical approach that supports people to record and store personally meaningful sounds, words, phrases, sentences, and/or stories over time, using their own natural voice, inflection, and intonation. These messages are selected to reflect the identity and personality of the speaker. If and when needed, the recorded messages can be incorporated into the most appropriate technology to enhance communication and support presentation of self (Costello, 2011, 2016). Because the emphasis is on the authenticity of the message rather than words, intonation, phrasing, and emotional content are important. There are no specific constraints on the quantity of messages needed as there is no requirement to deconstruct and reconstruct messages. Each recorded message can be stored in a unique location on an SGD. Although people who rely on AAC often depend on others to provide them with access to vocabulary and messages they may need to communicate (Beukelman & Light, 2020), message banking allows the person who uses AAC to control and direct vocabulary and message selection, resulting in a bespoke corpus that could never be curated by anyone else.

In the two decades since this clinical approach was first described (Costello, 2000), the term *message banking* has gained widespread currency, and the practice of recording messages has been introduced as part of clinical interventions in many different countries (e.g. see Oosthuizen, Dada, Bornman, & Koul, 2018), often through word-of-mouth or conference presentations. More recently, some companies have begun to use the term *message banking* to describe an option to record phrases directly into a specific device or software package.

The BCH message banking process: evolution and principles

The practice of recording messages may be implemented in many different ways and for a range of purposes. The sections that follow describe principles and steps associated with the message banking process at Boston Children's Hospital. The BCH process involves recording and storing messages long before a specific SGD or software option has been recommended, and in this sense it is technology-agnostic. More recently, the approach has evolved to incorporate "double dipping," which combines the potential of

message banking and voice banking to create personalized voices in a single integrated process.

Evolution of the BCH message banking process

Message banking was initially introduced as part of the Augmentative Communication Program at the Boston Children's Hospital to support patients facing significant speech changes due to medical interventions, enabling them to have a personal voice for relevant messages that reflected self (Costello, 2000). The goal was also for loved ones and medical care providers to hear the voice, personality, and emotion of the individual patient (Costello, 2009). Building upon the success of this in-patient program, in 2011 the BCH Message Banking Process was extended to include individuals with ALS who are at risk of permanently losing the ability to speak. In 2015, a formal dedicated ALS program was created. Since that time, more than 750 people have engaged with the program. The feedback, guidance, and outcomes from those individuals has directly informed the evolution of the BCH message banking process. The most recent innovation is the potential to use banked messages to create custom, synthetic voices through a process described as "double dipping," discussed later in this tutorial.

Principles of the BCH message banking process

Banking messages involves recording and capturing messages, phrases, words and sounds as one part of a multi-layered, dynamic, evolving process of engagement and support with individuals with neurodegenerative diseases (most often ALS) and with their loved ones. At BCH, a set of six guiding principles are at the core of the process: ownership and control, preserving identity, interpersonal and collaborative, dynamic and personalized, versatile and unique, and empowerment through hope. These principles have emerged through exploration of the literature, engagement with people with ALS, collaboration with clinical staff within the program, and detailed discussions during weekly reviews of patient assessments and outcomes. No principle is unique to message banking, but each has a distinct resonance in this context. All six are integral to the process, although the weight and emphasis may be dictated by a person's responsiveness and expressed investment, as well as individual clinical judgment at any point in a clinical interaction.

Ownership and control

The first principle reflects Article 3 of the Convention on the Rights of Persons with Disabilities (CRPD; United Nations, 2007) in respecting the autonomy and independence of all individuals, by investing ownership, power, control, and choice throughout the message banking process in individuals with ALS and their families. Many people with ALS report experiencing a loss of control over their outcome and their sense of self (Foley, Timonen, & Hardiman, 2014) and express a strong desire to control decisions related to their care (Regan, Preston, Eccles, & Simpson, 2019). In message

banking the aim is to give people power to maintain a piece of self and authority over how that self is projected in communication, even if they can no longer speak. Through providing information, the clinician aims to enable individuals with ALS and their families to make informed choices, recognizing that the extent of information needed or desired can vary greatly (Regan et al., 2019). Control extends to choices about whether to engage with message banking at all; how to engage; what to record; what messages to use in an SGD if needed; and when to stop recording, with no judgements about the content or extent of recorded messages. The only imposed structure is technical, to ensure that recordings are of sufficient quality for use in communication technology.

Preserving identity through voice

The second principle promotes the rights of individuals to express, construct, and preserve their personal and social identity through their unique acoustic fingerprint (Costello, 2000), as represented by their particular choice of words, phrasing, expressions, and vocal quality. The emphasis is on capturing *who you are now*, as manifested through spoken language. Reflecting on one's own communication style can be challenging. Concrete examples or questions (e.g. "What do you say to make your kids laugh and roll their eyes?" "Are you known for a catch-phrase at work?") may trigger specific recollections or interactions. Identifying personal examples of expressions of humor, sarcasm, or wit can illustrate the link between particular expressions and personal identity. Exploring "who you are now" may also open up emotionally challenging vistas of future uncertainties. Humor may ease some of this pain, but there is a need for continuous sensitivity to the in-the-moment emotional needs and capacities of individuals with ALS and their loved ones as they navigate this process. A recent clinical example of this was a 32-year-old man with ALS, whose wife was 7 months pregnant with their first child. While in the clinic, he chose to message bank many humorous animal sounds, pet names for his son "Hey there mini-me" and "I love you Champ." He began to quietly sob at the thought of his pending loss. After a few moments, he was offered a tissue with the comment: "I'm offering you the really good tissues from my private stock, not the sandpaper ones that the hospital offers." He laughed and then, while both crying and laughing, recorded a line from a song "It's my party and I'll cry if I want to" followed by the advice to his soon-to-be-born son, "Real men DO cry!"

Collaborative and relationship-focused

The third principle underpinning message banking is that it is an interpersonal, collaborative process, shared between people with ALS, their loved ones, and clinicians. Often, the sayings, mannerisms, and turns of phrase that mark the individual personality of a speaker are unconscious, even invisible to that speaker and may be most obvious to those close to them. Discussions on what to record can offer a rich opportunity to celebrate the uniqueness of each person with ALS. The experience of the AAC team at BCH is that the

most prolific recording outcomes are products of shared conversations with familiar partners. The collaboration extends to the relationship with the clinician. As has been reported with many other client groups, (e.g. people with aphasia; Hersh, Worrall, Howe, Sherratt, & Davidson, 2012), support is at its most effective when it is grounded in strong personal relationships with the person with ALS and with key communication partners. One example of this was Martha, who was observed throughout her appointment to make statements to her partner and son like, "You are so kind," "What would I do without you," and "Your talents can't be counted." As she reflected on messages to record, she listed many statements of need and self-advocacy. When the clinician pointed out his observation that she naturally related to other people by making them feel good about themselves, her family agreed that her style of social engagement was to lift up the people around her through her choice of language. Martha immediately stated "Oh my goodness, I was so focused on thinking about my needs, but you are all correct, I really connect with others by shining light on their strengths and gifts that I observe. I always want people to hear that, even if I can't speak."

A dynamic individualized process

The fourth principle of message banking is that the process is dynamic and evolving, extending from the initial point of contact, cycling through iterative phases in response to changing needs and opportunities. People living with ALS are constantly faced with the need to adapt and change as their disease progresses (Foley et al., 2014; King et al., 2008). The range of messages that could potentially be banked by an individual is unlimited; however, what is perceived as important to bank may change as a disease progresses. Initial messages may serve a primary role of archiving elements of a personal identity. Over time, individuals may decide to extend recordings to include messages associated with personal care, autonomy, medical decision making, and independence. Life changes, (e.g. the arrival of new family members) may motivate new recordings. A shift to using banked messages for interpersonal communication may occur at any point, or not at all. Each individual chooses what to record and how to use those recordings, and to adapt and change as often as desired over the course of the illness.

As an example, one woman began to bank messages shortly after her diagnosis when she had few symptoms. She banked many of her favorite quotes from literature, as well as statements in French and Latin that she liked to insert in conversation with her book club friends. Many months later, she banked messages that were germane to her new life experience (e.g. "Please, wipe my mouth"), conceding that she never conceived of these messages at the start, as they were foreign to her then-lived experience. She added that while her speech remained intelligible, she thought it important that people hear the sincere appreciation in her voice as she solicited assistance. She also requested that her story and her revelation be shared with those being introduced to the message banking process, so others could learn from her experience.

A unique and personal process

The fifth principle is that the process belongs to individuals and as such each iteration is unique, customized, and personal. Just as there is no one path to completing the process, there is no single outcome. All aspects are driven by the unique wishes, personality, needs, and aspirations of the individual who chooses to record and bank messages. Although there may be many shared experiences and challenges, the impact and meaning of those events are unique.

The importance of hope

The sixth principle is that message banking should offer each individual hope through empowerment, by ensuring from the outset that participating in this process is about taking control. It can be described as an insurance policy that may never be drawn down: Much like an insurance policy, its value lies in the reassurance that it is available if needed. Many people with ALS have commented on the value of this analogy. The process offers a way to make a positive investment in the future, regardless of the losses that may be brought on by the disease, offering an element of the hope that has been highlighted as a key coping mechanism (Freer, 2010).

Implementing the BCH process: a multi-dimensional personal process

Implementation of the six guiding principles requires careful planning and support over extended periods of time, as well as expertise and experience. Although from the point of view of the individual with ALS the most salient dimension may be the technicalities of recording and banking messages, successful implementation of BCH message banking requires a range of steps, supports, and resources that extend far beyond recording tools.

Key steps in implementing the BCH process

While individual clinical styles may vary, a structured 5-step process has been established based on successful patient outcomes (See Table 1). This process is a foundation of the orientation and training for clinicians in the ALS program. Individuals with ALS arriving at the clinic often report some awareness of the potential to "save a voice" but they are unclear about how this is accomplished. They may be interested in and even intrigued by the process but they and their loved ones may be emotionally paralyzed at the prospect of losing speech and voice, as well as what they perceive as an inevitably daunting task to preserve it. For this reason, it is critical that counseling and rapport building form the cornerstone of the message banking process.

Table 1. Key steps in implementing the BCH message banking process.

Steps	Practical resources and demonstrations
Build rapport and set the scene: Counsel in relation to impact of ALS; emphasize choice and control	Consider using analogy of insurance policy for message banking and/or concept of "wasting your time"
Invite person to articulate their goals; consider sharing personal view	Open questioning: What do you want us to accomplish today? What are you hoping we can achieve? What may I do to help you?
Provide information, introduce key terms	
Discuss concept of voice preservation and how strategies could be used in technologies	https://drive.google.com/file/d/1mxZWwE-LZE2rbrR4Ih_AQUL-MwLTedGp/view?usp=sharing
Introduce terms of message banking, voice banking, and double dipping	Demonstrate examples of voice banking and message banking embedded in communication tools (e.g. hand-held tablet, smartphone, speech generating device with alternative access)
Share clinical anecdotes and examples of messages recorded and used by others living with ALS	Share links to video resources, handout materials: www.youtube.com/watch?v=YlQp7rdku-s&t=1s
Demonstrate, practice, and check quality: Introduce the recording equipment and provide a tutorial on use	Provide hand-held recorder; review hardware and controls; provide recording guidance: (distance to hold, pressing controls);share online video tutorials: https://tinyurl.com/zoomhinoverview
Provide guidance for messages, (e.g. length and type of message) and record sample messages	Examples can include: (a) messages should not be so specific they can be used one time only, (b) record sentence starters such as "I need help with … ," (c) avoid recording names in messages for use with many people, (d) record a number of messages (repeat recording as necessary)
Provide guidance for voice banking	Provide headset and computer; create MMB account at Createmymessagebanking.com; share tutorial at www.youtube.com/watch?v=QQfd9BCG4pY&list=PLFWJQxERpRE-JStSZ4liJxeZMSM3SrXoV
Store and transcribe messages	Demonstrate process for uploading files from the recorder into MMB.com; demonstrate processes of transcription, editing, and tagging to message categories; review all features of the application (e.g. print, search)
Review possibility for double dipping and technical criteria	https://drive.google.com/file/d/1aFA9eN_UVfnS357wzC3fzFNobvbyKlbp/view?usp=sharing

Step 1. Building rapport and setting the scene

The introductory meeting about the prospect of voice preservation through voice banking and message banking typically lasts 90–120 min. The topic of banking messages often emerges from broader discussions about strategies to maintain existing function, the potential of assistive technology, and future concerns about communication. Banking messages involves discussing deeply personal and emotional topics, and so sufficient time and space to build rapport is crucial. Creating this rapport supports greater patient adherence and ultimately better outcomes (Ross, 2014). It is also associated with enhanced quality of life and patient satisfaction (Kornhaber et al., 2016).

The clinical introduction usually starts with the statement "You may not need any of the strategies I am going to describe. My hope is that this is a waste of your time" followed by introducing the analogy of the insurance policy, stressing that there is no predetermined course or timetable of any person's disease. To put the use of message banking in context, those involved are reassured that the goal is to partner with them to create options; although there is a possibility that banked messages or a custom synthetic voice may never be needed, a proactive plan is preferred over allowing a future need to emerge without effective preparation. This approach has elicited positive responses from people with ALS, (e.g. "So it is sort of like my home insurance policy; hopefully I'll not need to use it, but it is there if I do"; "It feels good to be able to do something that will positively impact my life when everything is taken away"). Often, patients have ended a clinical session commenting, "This has made me so hopeful" followed by the query "When can I schedule to come back so you can waste my time again?"

Step 2. Inviting the person to articulate their goal

Collaborative goal setting is at the core of patient-centered care and there is emerging evidence of its influence in intervention outcomes and therapeutic relationships (Vermunt, Harmsen, Westert, Rikkert, & Faber, 2017). Achieving effective collaboration is complex and can vary across individuals (Regan et al., 2019). The experience at BCH is that many individuals struggle to explain why they have come to their first appointment. Others express explicit goals ("I want to deliver the toast at my daughter's wedding"; "I need to make sure my grandchildren always hear ME state how much I adore them even if I can't talk"). One man highlighted "I want to make sure my children and grandchildren will still talk to me if I lose the ability to speak and that means I have to be able to talk with them." Another said he believed his voice was his signature and he wanted to preserve it, in contrast to another who stated "I'm just on a fact-finding mission."

Step 3. Introducing key terms

Prior to discussing specific terms, the potential use of AAC is discussed, highlighting distinctions between messages that might effectively be communicated using readily available synthetic speech options (e.g. "I need a drink") with more intimate or personal messages expressing affection, appreciation, or sarcasm, where a personal voice may be more important. This distinction is often introduced using personal reference aligned with the following script:

> If I ever had a difficult time speaking clearly, there are some messages that I would be fine communicating with an off-the-shelf synthetic voice. These include messages such as "open the door please" or "what time is it?" Modern synthetic speech is quite clear and of good quality. But there are many messages, regardless of the clarity of the synthetic speech, that I would want people around me to hear communicated exactly the way I

speak them. I want people to hear my intonation, with my emotion, with my heart, my passion, my appreciation, or even my sarcasm. I want people to hear ME authentically say messages like: "I am so proud of you" or "would you do me a favor, please?" or "I love you so much."

This discussion offers an opportunity to clarify that a back-up communication tool may only be needed at certain times for certain functions. It is intended to help people recognize that they do not need to record every word or sentence they may ever wish to speak. Above all, it is intended to help individuals to become "meta" about their own communication. Following this, several key terms – message banking, trademark messages, voice banking, and double dipping – are introduced and explained.

Message banking is described as a patient-driven strategy to support people who are at risk for changes in their speech to proactively record and store words, phrases, sentences, and personally meaningful sounds and/or stories, using their own natural voice, inflection, and intonation. If and when needed, these messages can be incorporated into an appropriate communication tool, identified through an evidence-based feature-match (Shane & Costello, 1994) assessment that addresses the motor, linguistic, physical access, and positioning needs of that person when natural speech becomes difficult to use. At that time, banked messages can be incorporated into the technology that best matches their identified needs and abilities.

Trademark messages are defined as those messages that are particular to an individual, often associated with an individual by their familiar communication partners ("Dad always says that to let you know he thinks you are wasting your money") and are typically delivered with unique intonation and prosody. People often quickly embrace this concept when encouraged to think of the messages as their "ism." During an introduction session, one teenage daughter replied to her father "Oh I can tell you all of your 'dad-isms'!" Messages may also include an individual's unique delivery of a message used by many other people; it does not need to be meaningful to the wider population but it may have a personal meaning to one person. Trademark messages do not have to be real words and may include varied types of laughter and sounds people may often make such as groans or sighs.

Voice banking is explained as a widely-used process of recording a large inventory of speech, based on a script designed to sample all co-articulated sounds in the target language. The recordings are used to create a synthetic voice that approximates the speaker's natural voice.

Double dipping is described as an option within the BCH model that involves using recorded messages that were collected through message banking, as the script for creating a personal synthetic voice. The option of using one effort for both message and voice banking, (i.e. double dipping), can eliminate the need to expend additional energy to complete each process separately.

Step 4. Demonstrating how to record and check messages

Once core terms have been explained, the discussion switches to practicalities. Even in a first visit, as people begin to identify messages they would like to bank, they often ask about the tools needed. The process of message banking can extend over many weeks and messages may be triggered by specific contexts, and so recording equipment must be constantly available and easy to access. The option of headset microphone input is important, to respond to changing motor abilities and enhance audio quality. Computer interface for file management should not be technically demanding, so that individuals can manage their own message files if they wish. As preservation and presentation of self are at the heart of message banking, high-quality audio files are essential: the 44/16 byte.WAV file format can be integrated with technologies ranging from iTunes to available communication software. Through trial with multiple tools, the Zoom H1and H1n recorders[1] have been identified as recorders that meet the necessary features. Some patients have expressed a wish to auto upload recorded files to Mymessgebanking.com. Other tools, such as smartphone technologies, potentially offer this option. To date, the experience at BCH has been that when smartphones have been used for recording, audio quality has varied. Sometimes quality has been acceptable for capturing banked messages but insufficient for double dipping, and some patients have had complications with file management. When recommended tools are not used there is a risk of poor outcomes, a need for multiple repetitions, loss of precious time, and potential loss of motivation and engagement.

Capturing spontaneity. Participants are counseled to carry the recorder with them throughout the day so that messages may be recorded spontaneously. Important messages may only become evident in context, such as at a grandchild's baseball game ("Way to go, that was great!") or when advocating for self ("I want you to explain everything to me before you speak with my wife"). Ideally, background noise should be minimal when recording but the importance of spontaneity often overrides this requirement. If a message is recorded in the moment, decisions can be made later about whether any need to be redone to maximize quality.

Technicalities of recording. Each person with ALS is given a handheld recorder to borrow for as long as needed. The recorder is set up in advance of the appointment and the operational features and potential problems (e.g. running out of recording space, changing the recording format, replacing the battery) are reviewed together at the first appointment. The mechanics of recording messages are demonstrated, highlighting the hardware, the sequence of button selection, the distance the recorder is held from the mouth, and the operational feedback of the recorder (i.e. the record light and the digital display readout on the recorder). The recorder is then given to the person with ALS to do a few practice recordings. It is often at this juncture that there is coaching from loved ones. If a person sounds too monotone, they are prompted to be more authentic; if they record naturally, they are congratulated. Family and friends often interject suggestions for what should be recorded, leading to laughter or moments of reminiscence. Patients are also sent

a link to video tutorials at: https://tinyurl.com/messagebankingprocessvideos (see also QR code 1, Appendix A, Supplemental online materials).

Quality check. Once a number of messages have been recorded, the process for downloading messages to a computer is demonstrated. Each message is played so the quality of recording can be reviewed. For individuals already experiencing changes in speech quality, this is often the first time they hear their affected voice and speech. Some simply comment on the changed quality and express an increased urgency to engage in message banking. Others are anxious to have a conversation about the deterioration in speech. It is critical that time is available to discuss each individual's unique perspective and concerns (Regan et al., 2019). Often family and loved ones highlight that while the speech may sound different, it remains the voice that they recognize and the voice they want preserved in the event of further speech changes.

Listening to the recordings together provides an opportunity to provide functional feedback and guidance. More coaching may be needed about positioning the recorder, pausing a moment prior to speaking, limiting the length of a message, and making sure background noise is minimal before recording. At this time, a handout is provided with many examples of messages banked by dozens of other people with ALS. Although not a script, the thousands of examples organized by generic categories such as greetings, physical needs, humor, sarcasm and self-advocacy may be useful tool to elicit ideas. The English version of the handout may be downloaded using QR code 2 (Appendix A, Supplemental materials).

Step 5. Transcribing and storing messages

Historically, the labor and time required to manage recorded files was overwhelming and often prohibitive. Files needed to be transferred, transcribed, saved, and categorized into related terms (e.g. personal needs) as a template for an SGD. If needed, page sets were created from these files within the communication software on an SGD. Then files were manually entered into the page set and the audio files were linked

to the text. This level of labor was extremely time-intensive and untenable for most clinicians.

To minimize much of the labor required to manage the audio files, a free web-based service that is not designed for any specific AAC technology (technology agnostic) was created in collaboration with TobiiDynavox[2] (www.mymessagebanking.com) to automatically create a text transcription for each sound file, eliminating many hours of manual work. Each message can now be tagged, automatically identifying it as belonging to an assigned category (or categories) such as humor, family, comfort, etc. The messages can be printed, for review if needed. This has been especially helpful for some people, one of whom commented "I can't believe I did not bank my wife's name! Thank goodness I have this printout so I can take inventory and make adjustments."

Several developers have modified their communication software to interface with the website, allowing for messages to be downloaded with pages and categories automatically created and populated with text and linked audio files. In a further recent innovation, one can now select the language of the recordings from a range of options, and the language in which the text transcription is to be provided (see Figure 1). This was achieved initially using IBM Watson[3] and then Microsoft Azure[4] to include more languages. For speakers with dysarthria, this transcription can be imprecise and require manual correction. While progress is being made with dysarthric speech recognition (e.g. the current research being led by Google through Project Euphonia), current commercial text-to-speech systems are designed for typical speakers (Young & Mihailidis, 2010) and further development is required for those with dysarthria (Schultz et al., 2021). After messages are uploaded, the website provides (a) a text transcription of the raw file to allow on-line editing, (b) a confidence level of the match of the transcription with the audio recording, (c) a playback option so the recording and transcription can be compared, and (d) the option to identify a category or a tag in which the message could be stored and found in an SGD (see Figure 2).

Many people with ALS prefer that file management be handled by the clinician, while others only ask the clinician to check messages and provide feedback; in order to facilitate this process, an account is created in mymessagebanking.com when everyone is in the clinic and the sign-on
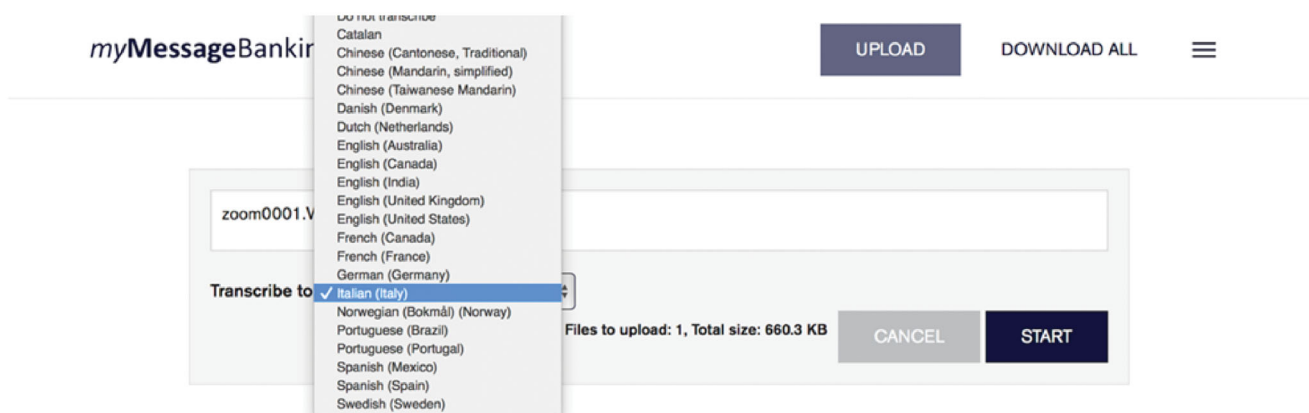

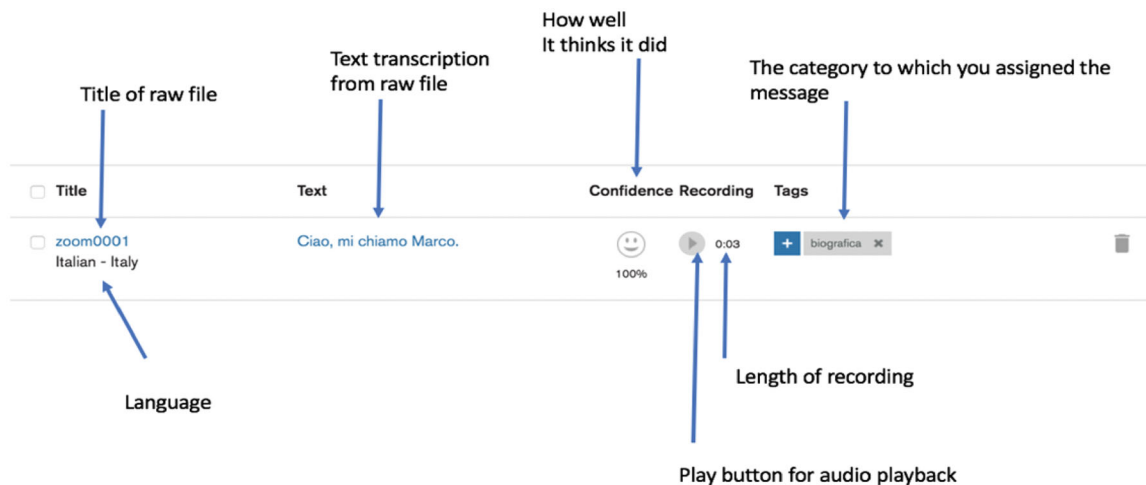**Figure 1.** Selecting language choice.

Figure 2. Functions for managing recordings.

credentials and password are shared with the clinician. This approach has the added benefit of facilitating distant support. The clinician can sign into the account, listen to recordings, check the transcription and tagging, and provide feedback, as needed.

Double dipping

When people embrace the concept of banking messages, many circle back to ask about voice banking. Creating a personalized synthetic voice allows someone to use text-to-speech to generate novel messages and to combine those with banked messages. It was experiences of people like "Laura" (described below), that led to the development of double dipping: using recordings completed through message banking as the data source for creating a synthetic voice.

> Laura, a woman with ALS, was reflecting on her unborn grandchild. She had already recorded "I love you Abigail" and "I love you Luca" for her two grandchildren as well as the general phrase "I love you." She wanted to be able to say, "I love you" and fill in the name of her yet-to-be-born grandchild using a voice that sounded like her. A quality voice created through voice banking afforded this flexibility. Voice banking options were reviewed and the process, time investment, and effort required for each was detailed. Examples of voices created with different methods were demonstrated. Laura reported that she did not wish to invest any time recording solely for voice banking. She valued every moment of life and preferred to spend her "healthy" time playing with her grandchildren.

At the time of writing, several voice banking companies had indicated that they are able to use the double dipping method to create a synthetic voice. One voice banking service with whom we have successfully exploited banked messages to create a personalized voice that patients have embraced, has been Acapela.[5] Undoubtedly others will become available and offer quality outcomes. As an example, using My-Own-Voice,[6] Acapela can take audio files downloaded from mymessagebanking.com and use those messages to create a synthetic voice. Acapela has successfully created dozens of high-quality, synthetic voices in multiple languages from banked messages. An example in Italian, comparing a banked message and the voice created through

double dipping, can be viewed at https://youtu.be/umGQZmvRSH8 or using QR code 3 in Appendix A (Supplemental materials). Successful double dipping requires access to a minimum of 550 but ideally 750 or more messages of sufficient quality. Clinical experience to date is that without exception, speakers have judged voices created through double dipping to more clearly approximate their biological voice compared to the voice created through voice banking alone.

Integrating message banking into AAC systems

When individuals and their families have discussed message banking, practiced with the recorder and listened to their messages, reviewed the web-based upload and transcription process, and explored the possibility of double-dipping, they often then ask, "What would I do with these if I ever needed to use them?" This offers a valuable opportunity to clarify the potential use of banked messages, as well as to initiate a wider discussion of SGD features and options. Since the launch of the service, hundreds of patients have successfully integrated recorded messages and the voice created through double dipping into their communication technology. Some have given permission to make video examples of messages and voice incorporated into AAC technology available at https://tinyurl.com/reallifeusemessagebanking (see also QR code 4, Appendix A, Supplemental materials).

Discussing how banked messages might be integrated into a wider communication system provides an opportunity to review software packages that support the audio file download and page creation. It offers a chance for people to become familiar with SGDs and access strategies such as eye-tracking, head mouse control, joystick control, and single/dual switch control. Exploring message and voice integration in various Windows, Android and iOS-based speech generating tools allows discussion of the wider AAC assessment process often led by patient questions, including requests to try specific pieces of technology. This discussion may not take place in the early stages, but the clinician must be mindful of the need to ensure that the opportunities and

limitations of message banking are considered in the context of an optimally flexible and effective communication system.

Implications for practice

The value of personal voice in AAC technology has gained much recognition in recent years, leading developers of communication software and SGDs to incorporate custom voice as an option into many devices. Commercial voice-creation companies historically have focused on creating a vocal identity for a corporation or other commercial products but have recognized the benefit of applying their expertise to the custom voice creation demand for AAC. At the same time, options for creating customized voices have expanded, with many premier companies dedicated to custom voice creation. These developments mean that the possibility of offering individuals who face potential loss of their natural speech their own custom voice is increasingly affordable, accessible, and implementable. With these new opportunities come responsibilities for clinicians to ensure that individuals are supported to maintain and preserve their personal identity as expressed through communication and, through this process, to reinforce their social roles and relationships.

A written clinical framework to support people at risk of losing the ability to speak has not existed until now. This tutorial may support clinicians in understanding the value of personal voice and recognizing the multiple options that should be thoughtfully introduced to a potential candidate for voice preservation. Throughout, emphasis has been placed on the importance of extending the concept of message banking beyond the mechanics of recording messages, through applying a framework that highlights the essential element of counseling and implementing the most appropriate voice preservation strategies in a personalized and dynamic approach. A further detailed illustration of the process is presented through the story of Maria (not her real name) and her engagement with the BCH Message Banking Process and Double Dipping as she navigated her journey with ALS (see Appendix B, Supplemental Material, online only).

Conclusion

Voice banking and the BCH Message Banking Process, including "double dipping" represent complementary approaches to maximizing communication for individuals faced with the potential loss of speech and associated potential loss of identity and social connectedness. Although the mechanics of this process are relatively simple, its success hinges on adherence to the principles discussed here and on meaningful relationships between all involved. The steps outlined guide clinicians in implementing this strategy, and the online resources provide further information.

Acknowledgements

Disclosure statement

Notes

1. Zoom North America, 2040 Express Dr S #500, Hauppauge, NY 11788.
2. Tobii Dynavox, Karlsrovägen 2 D SE-182 53 Danderyd, Sweden.
3. IBM Watson, 1 New Orchard Road, Armonk, New York 10504-1722.
4. Microsoft Azure, One Microsoft Way, Redmond, WA, USA.
5. Acapela Group, 33, Boulevard Dolez, 7000 Mons, Belgium. https://www.acapela-group.com
6. My Own Voice Acapela Group, Boulevard Dolez, 7000 Mons, Belgium. https://www.acapela-group.com

ORCID

John Costello 🆔 http://orcid.org/0000-0002-7010-4767
Martine Smith 🆔 http://orcid.org/0000-0003-2122-5607

References

Ball, L., Beukelman, D., & Pattee, G. (2002). Timing of speech deterioration in people with Amyotrophic Lateral Sclerosis. *Journal of Medical Speech-Language Pathology*, *10*(4), 231–235.

Beukelman, D., Fager, S., & Nordness, A. (2011). Communication support for people with ALS. *Neurology Research International*, *2011*, 714693. doi:10.1155/2011/714693

Beukelman, D., & Light, J. (2020). *Vocabulary selection and message management, augmentative and alternative communication: Supporting children and adults with complex communication needs* (5th ed.). Baltimore, MD: Brookes Publishers.

Bunnell, H. T., & Pennington, C. A. (2010). Advances in computer speech synthesis and implications for assistive technology. In J. Mullennix & S. Stern (Eds.), *Computer synthesized speech technologies: Tools for aiding impairment* (pp. 71–91). Hershey, PA: IGI Global.

Cave, R., & Bloch, S. (2021). Voice banking for people living with motor neurone disease: Views and expectations. *International Journal of Language & Communication Disorders*, *56*(1), 116–129. doi:10.1111/1460-6984.12588

Costello, J. (2000). AAC intervention in the intensive care unit: The Children's Hospital Boston model. *Augmentative and Alternative Communication*, *16*(3), 137–153. doi:10.1080/07434610012331279004

Costello, J. (2009). Last words, last connections: How augmentative communication can support children facing end of life. *The ASHA Leader*, *14*(16), 8–11. doi:10.1044/leader.FTR2.14162009

Costello, J. (2011). *Legacy messages: Message banking for people with amyotrophic lateral sclerosis and other acquired conditions*. San Diego, CA: American Speech Language Hearing Association.

Costello, J. (2016). Message banking vs. Voice banking: A very successful proactive model for people with ALS/MND. In *14th annual allied professionals forum*. Dublin, Ireland. https://www.alsmndalliance.org/wp-content/uploads/2016/12/costello-message-banking-handout.pdf

Foley, G., Timonen, V., & Hardiman, O. (2014). Exerting control and adapting to loss in amyotrophic lateral sclerosis. *Social Science & Medicine (1982)*, *101*, 113–119. doi:10.1016/j.socscimed.2013.11.003

Freer, S. (2010). Motor neurone disease: Insight into experience of family carers. *End of Life Care*, *4*, 54–61. doi:10.1136/EOLC-04-01.9

Fried-Oken, M., Mooney, A., & Peters, B. (2015). Supporting communication for patients with neurodegenerative disease. *NeuroRehabilitation*, *37*(1), 69–87. doi:1053-8135/15/$35.00

Google Research, Project Euphonia. https://sites.research.google/euphonia/about/

Hersh, D., Worrall, L., Howe, T., Sherratt, S., & Davidson, B. (2012). *SMARTER* goal setting in aphasia rehabilitatoin. *Aphasiology*, *26*(2), 220–233. doi:10.1080/02687038.640392

Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A-R., Jaitly, N., … Kingsbury, B. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, *29*(6), 82–97. doi:10.1109/MSP.2012.2205597

Juslin, P., Laukka, P., & Bänziger, T. (2018). The mirror to our soul? Comparisons of spontaneous and posed vocal expression of emotion. *Journal of Nonverbal Behavior*, *42*(1), 1–40. doi:10.1007/s10919-017-0268-x

King, S. J., Duke, M. M., & O'Connor, B. A. (2009). Living witih amytrophic lateral sclerosis/motor neurone disease (ALS/MND): Decision-makign about 'ongoing change and adaptation'. *Journal of Clinical Nursing*, *18*(5), 745–754. doi:10.1111/j.1365-2702.2008.02671.x

Kornhaber, R., Walsh, K., Duff, J., & Walker, K. (2016). Enhancing adult therapeutic interpersonal relationships in the acute health care setting: An integrative review. *Journal of Multidisciplinary Healthcare*, *9*, 537–546. doi:10.2147/JMDH.S116957

Maguinness, C., Roswandowitz, C., & von Kriegstein, K. (2018). Understanding the mechanisms of familar voice-identity recognition in the human brain. *Neuropsychologia*, *116*(Pt B), 179–193. doi:10.1016/j.neuropsychologia.2018.03.039

Makkonen, T., Ruottinen, H., Puhto, R., Helminen, M., & Palmio, J. (2018). Speech deterioration in amyotrophic lateral sclerosis (ALS) after manifestation of bulbar symptoms. *International Journal of Language & Communication Disorders*, *53*(2), 385–392. doi:10.1111/1460-6984.12357

McGettigan, C. (2015). The social life of voices: Studying the neural bases for the expression and perceptino of the self and others during spoken communication. *Frontiers in Human Neuroscience*, *9*, 129–124. doi:10.1080/07434618.2017.1342690

Nathanson, E. (2017). Native voice, self-concept and the moral case for personalized voice technology. *Disability and Rehabilitation*, *39*(1), 73–81. doi:10.3109/09638288.2016.1139193

Oosthuizen, I., Dada, S., Bornman, J., & Koul, R. (2018). Message banking: Perceptions of persons with motor neuron disease, significant others and clinicians. *International Journal of Speech-Language Pathology*, *20*(7), 756–765. doi:10.1080/17549507.2017.1356377

Pardo, J. (2012). Reflections on phonetic convergence: Speech perception does not mirror speech production. *Language and Linguistics Compass*, *6*(12), 753–767. doi:10.1002/lnc3.367

Pattee, G. L., Plowman, E. K., (Focht) Garand, K. L., Costello, J., Brooks, B. R., Berry, J. D., … Green, J. R. (2019). Provisional best practices guidelines for the evaluation of bulbar dysfunction in Amyotrophic Lateral Sclerosis. *Muscle & Nerve*, *59*(5), 531–536. doi:10.1002/mus.26408

Podesva, R., & Callier, P. (2015). Voice quality and identity. *Annual Review of Applied Linguistics*, *35*, 173–194. doi:10.1017/S0267190514000270

Pullin, G., Treviranus, J., Patel, R., & Higginbotham, J. (2017). Designing interaction, voice, and inclusion in AAC research. *Augmentative and Alternative Communication (Baltimore, Md. : 1985)*, *33*(3), 139–148. doi:10.1080/07434618.2017.1342690

Regan, L., Preston, N., Eccles, F., & Simpson, J. (2019). The views of adults with neurodegerative diseases on end-of-life care: A metasynthesis. *Aging & Mental Health*, *23*(2), 149–157. doi:10.1080/13607863.2017.1396578

Ross, L. (2014). Facilitating rapport through real patient encounters in health care professional education. *The Australasian Journal of Paramedicine*, *10*(4), 1–11. doi:10.33151/AJP.10.4.50

Schultz, B., Tarigoppula, V., Noffs, G., Rojas, S., van Er Walt, A., Grayden, D., & Vogel, A. (2021). Automatic speech recognition in neurodegenerative disease. *International Journal of Speech Technology*, *24*(3), 771–229. doi:10.1007/s10772-021-09836-w

Seltzer, L., Prososki, A., Ziegler, T., & Pollak, S. (2012). Instant messages vs. speech: Hormones and why we still need to hear each other. *Evolution and Human Behavior : Official Journal of the Human Behavior and Evolution Society*, *33*(1), 42–445. doi:10.1016/j.evolhumbehav.2011.05.004

Seltzer, L., Ziegler, T., & Pollak, S. (2010). Social vocalizations can release oxytocin in humans. *Proceedings. Biological Sciences*, *277*(1694), 2661–2666. doi:10.1098/rspb.2010.0567

Shane, H., & Costello, J. (1994, November). Augmentative communication assessment and the feature matching process. In Mini-seminar presented at the annual convention of the American Speech-Language-Hearing Association, New Orleans, LA.

Sidtis, D., & Kreiman, J. (2012). In the beginning was the familiar voice: Personally familiar voices in the evolutionary and contemporary biology of communication. *Integrative Psychological & Behavioral Science*, *46*(2), 146–159. doi:10.1007/s12124-011-9177-4

UN General Assembly, *Convention on the Rights of Persons with Disabilities : resolution / adopted by the General Assembly*, 24 January 2007, A/RES/61/106, available at: https://www.refworld.org/docid/45f973632.html [accessed 23 December 2021]

Vermunt, N., Harmsen, M., Westert, G., Rikkert, M., & Faber, M. (2017). Collaborative goal setting with elderly patients with chronic disease or multimorbidity: A systematic review. *BMC Geriatrics*, *17*(1), 167 doi:10.1186/s12877-017-0534-0

Young, V., & Mihailidis, A. (2010). Difficulties in automatic speech recognition of dysarthric speakers and implications for speech–based applications used by the elderly: a literature review. *Assistive technology: The official journal of RESNA*, *22*(2), 99–112. doi:10.1080/10400435.2010.483646