The treatment of apraxia of speech

Hurkmans, Josephus Johannes Stephanus

Chapter 6

Speech-Music Therapy for Aphasia (SMTA)

Speech and Music Therapy, an Innovative Joint Effort

S M T A

6.1 | History and definition

To date, all reported therapy programmes using musical elements in the treatment of AoS and aphasia, as described in the previous chapter, are either speech therapy programmes provided by a speech therapist without the contribution of a music therapist (as in MIT, Albert et al., 1973) or music therapy programmes without participation of a speech therapist (for example, in SIPARI, Jungblut et al., 2006). However, integrating speech and music therapy using all musical elements may produce a new therapeutic approach founded upon a combination of the strengths of both of these therapeutic approaches: the speech therapist's specific knowledge of neurological speech disorders and cueing strategies skills, and the music therapist's knowledge of musical parameters and specific skills of music composition. Therefore, a therapeutic approach combining elements of speech therapy and music therapy has been developed, which is called Speech-Music Therapy for Aphasia (SMTA, De Bruijn, Zielman, & Hurkmans, 2005). With the integration of a music therapist in SMTA, treatment effects may be maximised by adding more musical elements.

Speech-Music Therapy for Aphasia (SMTA, De Bruijn et al., 2005) is a treatment programme for AoS and aphasia in the rate and rhythm control strategy category (see Chapter 3). It is a combination of speech therapy and music therapy in which a speech therapist and a music therapist provide the therapy *simultaneously*. The fundamental ideas for SMTA are based on various similarities between language and music, as described in Chapter 4.

SMTA is designed for patients with AoS and aphasia. Two patient groups can be distinguished: (1) non-speaking patients; and (2) non-fluent speaking patients. These two target groups will be described below.

6.2 | Target groups

6.2.1 | Non-speaking patients

Non-speaking patients are the most severe AoS patients with a total inability to articulate, even a single phoneme (mutism). These patients suffer from a severe disorder in initiating speech production. Also, patients with a severe aphasia, such as global aphasia, are classified in this category.

6.2.2 | Non-fluent speaking patients

Non-fluent speaking AoS patients suffer from prosodic impairments, which cause disturbances in the flow and melody of speech production. The prosodic impairments result in false starts, repairs, pauses and repetitive attempts at initiating speech production. Furthermore, patients in this category may show segmental impairments, such as phonetic distortions, causing disturbances in word accuracy and impairments in consistency (see Chapter 2 for an elaborate description of the AoS symptoms). In clinical practice, most patients with AoS also experience some degree of non-fluent aphasia, such as Broca's aphasia and conduction aphasia.

The above-described patient groups would imply that SMTA is not indicated for fluent-speaking patients, such as in Wernicke's aphasia. However, these patients may show AoS symptoms such as groping, repairs and phonetic distortions, as well. SMTA is then indicated to reduce these symptoms.

6.3 | Objectives

For each patient, the aims are personalised, dependent on the severity and symptoms of the speech disorder, and the needs of the patient. The *general* aims of SMTA are different for non-speaking and non-fluent speaking patients. First, for non-speaking patients, SMTA treatment

is aimed at 'de-blocking'. From the early days of aphasia rehabilitation, it is known that mute patients may be able to use language when singing. Various researchers found a preservation of singing in non-speaking aphasic patients (e.g., Broca, 1861; Falret, 1867). This finding has been reported in several later cases (e.g., Yamadori, Osumi, Masuhara, & Okubo, 1977; Hébert, Racette, Gagnon, & Peretz, 2003; Straube, Schulz, Geipel, Mentzel, & Miltner, 2008).

Second, for non-fluent speaking patients, SMTA treatment is aimed at improving speech motor programming and planning. This means improvement in: (1) accuracy; (2) consistency; and (3) fluency (i.e., the flow and melody of speech) of articulation.

6.4 | Treatment methodology

SMTA consists of two interwoven lines of treatment: (1) the speech-therapy line, and (2) the music-therapy line. Although the speech therapist and the music therapist work together with the AoS patient at the same time and in the same room, the two lines of treatment will be described separately.

Each SMTA session starts at the phoneme level with single phonemes and/or phoneme sequences to 'warm up' the voice. Depending on the target objectives, the exercises build up from the phoneme level towards the word and sentence level. The speech therapist selects the target items. Subsequently, the music therapist creates a composition of these items including various musical elements, usually played on a piano or a guitar.

6.4.1 | The speech-therapy line of treatment

The speech-therapy line of treatment consists of three levels: (1) phonemes (including syllables); (2) words; and (3) sentences. Exercises at the phoneme-level target improvement in the production of phonemes and syllables. For phonemes, vowels can be trained in isolation.

Usually sequences of three vowels are used. These can be sequential, for example, "aa" – "aa" – "aa", or alternating, for example, "aa" – "oo" – "ee". Phoneme selection in the sequence is based on distinctive features of articulation. Usually consonants are not practiced in isolation because they may obstruct the airflow and, therefore, limit the speech fluency, but they are used in syllables. As with vowels, syllables may be trained sequentially, for example "maa" – "maa" – "maa", or in an alternating manner, such as, "maa" – "moo" – "mee". At this level, initial clusters can be used, for example, "smaa" – "smoo" – "smee". Trained items at the phoneme level are *always* related to the target items at the word level. For example, "fra" can be trained at the phoneme level in preparation to the word "Frans", a common Dutch name, at the word level.

The word-level exercises are designed to be functionally relevant for AoS patients. Names of family members, places holding meaning, and other words deemed important to the AoS patients are trained. These personally relevant words are practiced alongside frequently occurring phrases such as "hello" (i.e., formulaic language). At the word level, common daily utterances, such as "good morning", are trained as well. These include more than one word but are not considered a sentence.

At the sentence level, selected sentences that are functionally relevant to individuals with AoS are trained. For example, "enjoy your meal" and "can you help me?" In selecting target sentences, grammar is inferior to comprehensibility. For example, at a restaurant the phrase "Inside or outside?" is preferred instead of "Would you prefer sitting inside or outside?"

The speech therapist uses various cueing strategies, including (1) phonetic cueing (i.e., auditory presentation of the first phoneme of a syllable, word or sentence), (2) visual cueing (i.e., showing mouth references), and (3) gestures (i.e., natural gestures supporting daily utterances, such as waving the hand while saying 'goodbye').

6.4.2 | The music therapy line of treatment

The musical line of treatment follows a structured procedure from singing to rhythmical chanting and speaking. Each and every selected target item follows this structure. The musical interventions are designed to musically support the speech exercises, and thus, they share the same structural composition from phoneme level to word and sentence level (De Bruijn, Hurkmans, & Zielman, 2011).

At the phoneme level, the music therapist uses scales or parts of scales because they are easy and known to the individual with AoS and aphasia. At the word and sentence level, the music therapist composes *new* melodies. SMTA does not make use of familiar songs because the language output in familiar songs is usually automatically generated (e.g., Schön, Gordon, & Besson, 2005).

Using tempo, meter, rhythm and dynamic parameters, the music therapist is able to support the melody in order to closely follow the prosodic features of speech production. These various musical parameters will be described below and originate from De Bruijn, et al. (2011).

6.4.2.1 | *Melody*

Melody relates to pitch. Various pitches assemble in musical scales and these are frequently used in MT, because their structure is familiar to the patient (De Bruijn et al., 2011). Parts of the scales may be suitable such as, the first three, four or five tones of the scale, as they are short and they represent limited pitch ranges. Scales determine the exercise's level of difficulty and should match the patient's level of concentration. For example, the singing of strings of clusters (e.g., "stra"-"spra") requires effort. By limiting the number of tones sung in the scale, the therapist may prevent errors in articulation.

When practicing single words, the music therapist composes new melodies that best represent normal intonation and articulation. Musical

variations that deviate from these patterns are not recommended, for example, lengthening notes to emphasise endings or using syncopations to make the melody 'swing', as these distract the patient. To ensure that the syllable structure remains intact, it is imperative that each syllable is assigned only to one tone. For example, "Ju-dith" has to be composed at two notes and not at three resulting in 'Ju-hu-dith'. In songs, this is accepted (such as in Baroque music) but not in speech production. However, in words with a medial cluster, it may be necessary to insert a short pause within the word: "sister" → "sis- ter". An upbeat can be used to divide two consonants in a word with an initial cluster. For instance, when the word "playing" is too difficult, an upbeat is inserted resulting in a schwa-sound: "pu-lay-ing". After the cluster is well articulated using this intervention, removing the upbeat may reconnect the consonants: "play-ing". Finally, a word that is practiced in earlier sessions may be elicited merely by playing its associated melody.

To ensure that the patient associates a melody with a target item, a new melody needs to be composed for each new target item. Familiar tunes and melodies introduced earlier cannot be used in case they automatically evoke the recall of other utterances (Peretz, Gagnon, Hébert, & Macoir, 2004; Schön et al., 2005; Stahl, Kotz, Henseler, Turner, & Geyer, 2011). Each newly composed melody should be clearly distinguishable from all others as it is the melody itself that elicits the production of the utterance.

As mentioned earlier, the transition from singing to speech is only possible when the melodies closely match the prosodic features of speech production. The use of melisma, syncopation, accidental notes, complex melodic structures or large melodic intervals should, therefore, be avoided as these are not features of prosody and intonation in the Western language. The music therapist should compose simple melodies and ignore his or her desire for 'musical' creativity or excellence when composing melodies. This is particularly important at the begin-

ning of the therapy period when the structure of the melodies should be clear. The melodies should be easy to listen to including an adequate pitch range that matches the vocal range of the patient. Melodies that are too high or too low in pitch, limit the efficacy of the exercise, while complicated melodies may distract the patient.

Musical variation can be introduced when fluency of articulation improves. For example, tonality (major/minor) can be introduced to enhance the expression of various emotions. Certain common intervals may be used, such as the 'ascending fourth' as a signal for attention, for example in "Watch out!"

6.4.2.2 | Rhythm

Rhythm relates to duration: long and short. Some characteristics of rhythm are important in MT: the order of long and short note-values influences the degree of rhythmic complexity and, therefore, the level of difficulty of the exercises. For example, in 4/4 time the sequence long-short-short is more difficult than short-short-long because it allows the patient less time to prepare himself for the repetition of the exercise. Furthermore, syncopation, which is the irregular flow of rhythm, is not part of natural speech, and should therefore be avoided in MT.

The prosodic features of the target item determine the selected rhythm by the music therapist. For example, the daily utterance "good morning" should have the rhythm short-long-short because the accent is located at the second syllable, and, therefore, this syllable should be long.

6.4.2.3 | Meter

Meter directly relates to rhythm: the duration of long and short has to be embedded in a frame; a beat. Some distinctive features of meter are the following: 4/4 and 2/4 beats are supportive, familiar, and easy to sing; 3/4 and 3/8 beats evoke a swaying motion, and are suitable when

relaxation (decrease in tension) should enhance the patient's performance in the exercises. Finally, the movement stimulated by the 6/8 beat is relaxing, but may lead to an increase in tempo.

6.4.2.4 | *Tempo*

MT considers tempo to be the key for melodic adaptation. Some features of tempo are well known in MT. A slow tempo, for example, creates a sense of relaxation. However, a tempo that is too slow becomes static: there is no flowing motion. A tempo that is too slow does not stimulate the patient and may negatively affect the patient's singing. Contrarily, a fast tempo may stimulate the patient, and increase the level of concentration required to perform the task. When the tempo is too fast, the patient will not be able to articulate accurately. Therefore, when a new target item is introduced, the tempo usually is slow. While word accuracy improves, the tempo can increase, and, thereby, increasing the difficulty of the exercise.

6.4.2.5 | *Dynamics*

Dynamics relates to volume. In MT, dynamics ranges from mezzo-piano to mezzo-forte, which is usually the least taxing on the voice. However, the use of forte or even fortissimo can be necessary, for example to call or warn someone or to express emotions by the use of emotional prosody. For example, when the sentence "Peter, watch out!" is targeted, the top of a patient's volume is required.