IT HAS BEEN REPORTED THAT PATIENTS WITH SEVERELY nonfluent aphasia are better at singing lyrics than speaking the same words. This observation inspired the development of Melodic Intonation Therapy (MIT), a treatment whose effects have been shown, but whose efficacy is unproven and neural correlates remain unidentified. Because of its potential to engage/unmask language-capable regions in the unaffected right hemisphere, MIT is particularly well suited for patients with large left-hemisphere lesions. Using two patients with similar impairments and stroke size/location, we show the effects of MIT and a control intervention. Both interventions’ post-treatment outcomes revealed significant improvement in propositional speech that generalized to unpracticed words and phrases; however, the MIT-treated patient’s gains surpassed those of the control-treated patient. Treatment-associated imaging changes indicate that MIT’s unique engagement of the right hemisphere, both through singing and tapping with the left hand to prime the sensorimotor and premotor cortices for articulation, accounts for its effect over nonintoned speech therapy.

Received September 12, 2007, accepted January 5, 2008.

Key words: Melodic Intonation Therapy, aphasia, language recovery, brain plasticity, music therapy

O F THE ESTIMATED 600,000-750,000 NEW STROKES occurring in the US each year (according to data presented by the American Heart Association and the National Institutes of Health), approximately 20% result in some form of aphasia (http://www.wrongdiagnosis.com/artic/ninds_aphasia_information_page_ninds.htm). Aphasia is a condition characterized by either partial or total loss of the ability to communicate verbally. A person with aphasia may have difficulty speaking, reading, writing, recognizing the names of objects, and/or understanding what other people have said. Aphasia, a disorder caused by a brain injury (e.g., stroke, tumor, or trauma), can be subdivided into fluent and nonfluent categories. Nonfluent aphasia (as in the patients to be discussed here) generally results from lesions in the frontal lobe including the portion of the left frontal lobe known as Broca’s region. Named for Paul Broca (1864), who first associated this area of the brain with nonfluent aphasia, this region is thought to consist of the posterior inferior frontal gyrus (IFG) encompassing Brodmann’s areas 44 and 45. However, subsequent reports have shown that a wider array of lesions in the frontal lobes and in subcortical brain structures can also present a clinical picture of a Broca’s aphasia (see Kertesz, Lesk, & McCabe, 1977).

Surprisingly, there are no universally accepted methods for the treatment of nonfluent aphasia against which new or existing interventions can be tested, nor have any criteria been established for determining treatment efficacy. Most interventions in the subacute phase are conducted by speech therapists who evaluate patients’ individual needs, then use a combination of techniques to help recover language/facilitate communication. Despite the lack of specific criterion for success, most therapists would agree that treatment efficacy would be defined by patients’ ability to show improvement in speech output that generalizes to untrained language structures and/or contexts (Thompson & Shapiro, 2007).

Because the neural processes that underlie post-stroke language recovery remain largely unknown, it has not been possible to effectively target them using specific therapies. To date, functional imaging (mostly positron emission tomography) of language recovery has largely focused on spontaneous recovery, and patients have been imaged only after natural recovery has run its course (Warburton, Price, Swinburn, & Price, 1999; Weiller et al., 1995). Some studies emphasize the role of preserved language function in the left hemisphere (Cappa & Vallar, 1992; Heiss, Kessler, Thiel, Ghaemi, & Karbe, 1999), while others propose that language function is restored when right-hemisphere regions compensate for the loss (Basso, 1989; Blasi, Young, Tansy, Petersen, Snyder, & Corbetta, 2002; Cappa & Vallar, 1992; Cappa et al., 1997; Kinsbourne, 1998; Moore, 1989;
Selnes, 1999; Weiller et al., 1995). Still other studies report evidence for bihemispheric language processing (Heiss & Thiel, 2006; Mimura, Kato, Sano, Kojima, Naeser, & Kashima, 1998; Rosen et al., 2000; Saur et al., 2006; Winhuisen et al., 2005). Interestingly, only a few studies have examined the neural correlates of an aphasiasia treatment by contrasting pre- and post-therapy assessments (Cornelissen, Laine, Tarkiainen, Järvensivu, Martin, & Salmelin, 2003; Musso, Weiller, Kiebel, Muller, Bulau, & Rijntjes, 1999; Saur et al., 2006; Small, Flores, & Noll, 1998; Thompson & Shapiro, 2005). The general consensus is that there are two routes to recovery. In patients with small lesions, there tends to be more activation of left hemisphere peri-lesional cortex and variable right hemisphere activation during the recovery process or after recovery. In patients with large left-hemisphere lesions that involve language regions in the frontotemporal lobes, there tends to be more activation of the homologous language-capable regions in the right hemisphere.

Assuming that potential facilitators of language recovery may be either undamaged portions of the left-hemisphere language network, language-capable regions in the right hemisphere, or both, it is necessary to explore treatments that can better engage these regions and ultimately, change the course of natural recovery through neural reorganization. One therapy capable of engaging regions in both hemispheres is Melodic Intonation Therapy (MIT; Albert, Sparks, & Helm, 1973; Sparks, Helm, & Albert, 1974), a method developed in response to the observation that severely aphasic patients can often produce well-articulated, linguistically accurate words while singing, but not during speech (Gerstman, 1964; Geschwind, 1971; Hebert, Racette, Gagnon, & Peretz, 2003; Keith & Aronson, 1975; Kinsella, Prior, & Murray, 1988). MIT is a hierarchically structured treatment that uses intoned (sung) patterns to exaggerate the normal melodic content of speech at three levels of difficulty. The intonation works by translating prosodic speech patterns (sung phrases) using just two pitches. The higher pitches represent the syllables that would naturally be stressed (accented) during speech (see Figure 1). At the simplest level, patients learn to intone (sing) a series of 2-syllable words/phrases (e.g., “Water,” “Ice cream,” “Bathroom”) or simple, 2- or 3-syllable social phrases (e.g., “Thank you,” “I love you”). As each level is mastered, patients move to the next, and phrases gradually increase in length (e.g., “I am thirsty,” “A cup of coffee, please”). Beyond the increased phrase length, the primary change from level to level of MIT lies in the way the treatment is administered and the degree of support that is provided by the therapist.

MIT contains two unique elements that set it apart from other, non-intonation-based therapies: (1) the melodic intonation (singing) with its inherent continuous voicing, and (2) the rhythmic tapping of each syllable (using the patient’s left hand) while phrases are intoned and repeated. Since the initial account of its successful use in three chronic, nonfluent (Broca’s) aphasics (Albert, Sparks, & Helm, 1973), reports have outlined a comprehensive program of MIT (Helm-Estabrooks & Albert, 1991; Sparks & Holland, 1976) including strict patient selection criteria (Helm-Estabrooks, Nicholas, & Morgan, 1989), and data that showed significant improvement on the Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983) after treatment (Bonakdarpour, Eftekharzadeh, & Ashayeri, 2000; Sparks, Helm, & Albert, 1974). In a case study comparing MIT to a non-melodic control therapy, Wilson, Parsons, and Reutens (2006) found that MIT had a general facilitating effect on articulation, and a longer-term effect on phrase production that they attributed specifically to its melodic component. However, the outcomes of that study were measured by the patient’s ability to produce practiced phrases prompted by the therapist, rather than by the transfer of language skills to untrained structures and/or contexts.

Another important characteristic of MIT is that, unlike many therapies administered in the chronic phase that involve one to two short sessions per week, MIT engages
patients in intensive treatment totaling 1.5 hrs/day, five
days/week until the patient has mastered all three levels
of MIT. In addition to its unique elements, there are several
other components that play an important role in MIT, but
are also used by other therapies, among them are the slow
rate of vocalization (one syllable/s) and an administration
protocol that includes one-on-one sessions with a therapi-
list who introduces and practices words/phrases using
picture cues while giving continuous feedback. These
shared features were carefully considered as we designed a
control intervention for MIT that included the elements
common to other therapies while specifically excluding the
melodic intonation/continuous voicing and rhythmic tap-
ing that may likely be the key factors in its effectiveness.

The original interpretation of MIT’s path to success-
ful recovery was that it engaged expressive language areas in the right hemisphere (Albert et al., 1973; Sparks et al., 1974), although to date, this has not yet been proven. Alternatively MIT may exert its effect by either unmasking existing music/language connections in both hemispheres, or by engaging preserved language-
capable regions in either or both hemispheres. Since
MIT incorporates both melodic and rhythmic aspects of music (Albert et al., 1973; Boucher, Garcia, Fleurant, & Paradis, 2001; Cohen & Masse, 1993; Helm-Estabrooks & Albert, 1991; Sparks & Holland, 1976), it may be
unique in its potential ability to engage both hemi-
ospheres. Belin et al. (1996) suggested that MIT-facili-
tated recovery was associated with the reactivation of
left-hemisphere regions, most notably the left pre-
frontal cortex, just anterior to Broca’s region. Although,
this publication was the first to examine patients treated
with an MIT-like intervention using functional neuro-
imaging, their findings were both surprising and somewhat contrary to the hypotheses proposed by the developers of MIT (Albert et al., 1973; Sparks et al., 1974). Furthermore, to help interpret Belin and col-
leagues’ findings it is important to consider the follow-
ing: First, only two of their seven patients had Broca’s
aphasia; the rest were diagnosed with global aphasia.
Second, they conducted only one imaging session,
which took place after therapy (no pre-/ post- comparison).
And finally, their analysis was done in predefined
regions of interest rather than across the entire brain
space. It is interesting to note that although Belin and
colleagues’ primary finding was an activation of left
prefrontal regions when participants were asked to
repeat intoned words, there is an important aspect of
their study that is not often reported. In their analysis
comparing the repetition of spoken words with the
hearing of those words, they found blood flow changes
that occurred predominantly in the right hemisphere
(including the right temporal lobe and the right central
operculum), which concurs with some of our findings
detailed below.

The aim of our study is to describe and discuss the
unique and shared elements of MIT and to contrast the
behavioral and neural treatment effects of MIT with a
control intervention, Speech Repetition Therapy (SRT),
in two prototypical patients.

Materials and Methods

Participants

We present two of the patients we have treated to date in
our ongoing study of MIT’s effects. Both were diag-
nosed with severe nonfluent aphasia (restricted verbal
output, impaired naming and repetition, relatively
unimpaired comprehension) as the result of a left-
hemisphere ischemic stroke involving mainly the supe-
rior division of the middle cerebral artery, and classified
as having Broca’s aphasia. Despite the fact that the
patients had already received more than one year of tra-
ditional speech therapy prior to enrollment in our
study, they presented with significantly impaired verbal
output and remained unable to speak fluently. Both
patients were tested twice prior to therapy to establish a
stable baseline. In addition, we assessed their ability to
speak/sing the lyrics of familiar songs by analyzing the
number of Correct Information Units (CIUs; correct,
meaningful words) that each patient produced while
singing and speaking compared to the total number of
words for at least two familiar songs. Patient #1 (male;
age 47; right-handed; native language English; 12+
years of schooling; 2-3 years of instrumental practice as
a child; no active singing in a choir; moderate to severe
right hemiparesis; independent in activities of daily liv-
ing, ADL; Barthel-index of 95 out of 100) underwent an
intensive course of MIT and was assessed on behavioral
and neural measures at a series of regular intervals that
included assessments after 40 and 75 sessions of MIT.
Patient #2, (male; age 58; right-handed; native language
English; 12+ years of schooling; 2-3 years of instrumen-
tal practice as a child and some singing in choirs in high
school and college; moderate to severe right hemipare-
sis; independent in activities of daily living, ADL; Bar-
thel-Index of 95 out of 100), matched to Patient #1
with regard to lesion size/location and baseline speech
production abilities, underwent an equally intensive
alternative intervention, SRT, designed to control for
the elements of MIT that are common to other speech
therapies, and exclude its distinct features, the melodic
intonation and rhythmic tapping with the left hand.
After undergoing 40 sessions of SRT and the same series of behavioral and neural assessments administered to Patient #1, Patient #2 underwent treatment with MIT and was assessed after 40 and 75 therapy sessions. Thus, both patients had behavioral and brain imaging assessments before and after therapy. The proportion of spoken CIUs/total words possible was significantly lower than the proportion of sung CIUs/total words in both patients. Although both patients actually received 75 sessions of MIT, the comparison between the two interventions reported here was made after each patient had received 40 sessions of their originally assigned treatment (MIT and SRT respectively).

Language Assessments

Based on their Boston Diagnostic Aphasia Examination (BDAE; Goodglass & Kaplan, 1983) scores, both patients were classified as having Broca’s aphasia. Patients underwent a series of language assessments (see below) at baseline, and again, four weeks later. This was done in order to establish a stable baseline and record any fluctuations in performance. The same set of tests was administered to both patients after 40 treatment sessions (post40). Further assessments were done after 75 sessions (post75). The test battery consisted of the following speech production measures designed to quantitatively assess spontaneous speech: (1) Conversational interview: regarding patients’ biographical data, medical history, post-stroke treatment, daily activities, etc.; and (2) Descriptions of complex pictures: using patients’ responses on these measures, we calculated the average number of CIUs/min and the average number of syllables/phrase. All meaningless utterances, inappropriate exclamations, incorrect responses (inaccurate information), and/or perseverations were excluded prior to scoring. Participants were also given confrontation picture naming tasks, including the Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 2001) and a matched subset (30 images) of the Snodgrass-Vanderwart color pictures (1980). All behavioral assessments were videotaped for analysis. Videotapes were transcribed and patients’ speech output was checked for intelligibility, then scored by an independent researcher who was not associated with the patients during therapy.

Experimental Stimuli and fMRI Paradigm

A list of 16 bisyllabic words/phrases that both patients were capable of saying at baseline were used for stimuli in the fMRI experimental task at all imaging time points, and the rate of speaking/singing (one syllable/s) remained constant throughout the study. The functional task consisted of five conditions: two experimental (spoken or sung bisyllabic words/phrases) and three control (humming, phonation, and silence). In the experimental conditions, participants heard an investigator saying/singing two-syllable words or phrases (presented at the rate of one syllable/s), then repeated exactly what they had heard after an auditory cue. In the silence condition, participants were asked to wait for the cue, then take a breath as if to respond as they had in the other conditions (for more details on the fMRI tasks see Ozdemir, Norton, & Schlaug, 2006). Functional magnetic resonance imaging (fMRI) and sparse temporal sampling was done as previously reported in detail (see Gaab, Gaser, Zaehele, Chen, & Schlaug, 2003, and Ozdemir, Norton, & Schlaug, 2006). Participants’ responses were recorded for offline analysis. Both of the patients reported here had 100% correct response rates.

Treatments

The two patients in this study were randomly assigned to treatment type (either MIT or SRT). Both patients worked one-on-one with the same therapist for 1.5 hours/day, five days/week, and were given a set of materials for daily home practice. Both interventions were identical with regard to the length of phrases, the use of picture stimuli, and the level of support provided by the therapist at each stage of advancement. SRT differed only in that the phrases were spoken rather than intoned (sung), syllables were not sustained, and there was no hand tapping associated with the production of speech.

Results

Behavioral and Imaging Effects of the MIT Intervention

Patient #1, who was 13 months post onset of a left-hemispheric stroke (see Figure 2 for lesion location/size), was assigned to treatment with MIT. He underwent two pre-treatment assessments separated by 4 weeks (pre1 and pre2), a mid-treatment assessment after 40 therapy sessions (post40), and a post-treatment assessment after 75 sessions of MIT (post75). At baseline, his spontaneous speech assessments yielded results consistent with his diagnosis. Repeat assessments conducted prior to MIT showed no significant changes. After only 40 sessions of MIT, he showed significant improvement on measures of speech output and confrontational naming, and after 75 sessions of MIT, those improvements were even more pronounced (for all behavioral results, see Table 1).
Patient #1’s baseline and post40 fMRI studies (see Figure 3 in color plate section) showed posterior perisylvian activation on the left, and both superior temporal and inferior precentral gyrus activation on the right during the speaking condition (speaking vs. silence contrast); however, the post40 scan also showed more prominent right-hemispheric activation involving the right posterior middle premotor cortex and right inferior frontal gyrus, as well as a slightly smaller increase in activation of the posterior superior temporal gyrus.

 Behavioral and Imaging Effects of the SRT Intervention

Patient #2, who was 12 months post onset of a left-hemispheric stroke (see Figure 2 for lesion location/size), was assigned to treatment with SRT. His dissociation between speaking and singing and baseline speech production rate were similar to that of Patient #1. There were no significant changes on repeat baseline assessments. After 40 SRT sessions, Patient #2’s speech production scores improved, his picture-naming score increased (Table 1), and his fMRI studies (Figure 3 in color plate section) that had at baseline shown activation of the posterior superior temporal gyrus (STG), superior temporal sulcus (STS), and middle to inferior precentral gyrus on the right, with a very small area of activation on the left during the speaking condition (Overt Speaking vs. Silence control condition), showed more prominent left-hemispheric activation involving the inferior part of the pre- and post-central gyrus, as well as the middle and posterior portions of the STG/STS, left more than right. Patient #2 also shows the Overt Speaking vs. Silence (control condition) contrast after an additional 40 sessions of MIT that followed the SRT sessions. Slight differences can be seen in the regional magnitude of activation, with a greater emphasis on the right premotor/motor and temporal lobes.

TABLE 1. Summary of Language Outcomes.

<table>
<thead>
<tr>
<th>ID</th>
<th>Treatment</th>
<th>Measure</th>
<th>Baseline</th>
<th>Post40</th>
<th>% Change</th>
<th>Post75</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>MIT</td>
<td>CIUs/min.</td>
<td>4.40</td>
<td>10.10</td>
<td>229.50</td>
<td>13.90</td>
<td>315.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Syllables/phrase</td>
<td>1.80</td>
<td>4.10</td>
<td>227.80</td>
<td>4.70</td>
<td>261.10</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Picture naming (% correct)</td>
<td>60.00</td>
<td>80.00</td>
<td>133.30</td>
<td>95.00</td>
<td>158.30</td>
</tr>
<tr>
<td>Patient 2</td>
<td>SRT</td>
<td>CIUs/min.</td>
<td>3.60</td>
<td>6.80</td>
<td>188.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Syllables/phrase</td>
<td>2.40</td>
<td>4.00</td>
<td>166.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Picture naming (% correct)</td>
<td>59.00</td>
<td>72.00</td>
<td>122.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient 2</td>
<td>MIT (after SRT)</td>
<td>CIUs/min.</td>
<td>6.80</td>
<td>16.70</td>
<td>245.60</td>
<td>20.50</td>
<td>301.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Syllables/phrase</td>
<td>4.00</td>
<td>8.90</td>
<td>222.50</td>
<td>10.10</td>
<td>252.50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Picture naming (% correct)</td>
<td>72.00</td>
<td>90.00</td>
<td>125.00</td>
<td>89.00</td>
<td>123.60</td>
</tr>
</tbody>
</table>

Note: MIT = Melodic Intonation Therapy; SRT = Speech Repetition Therapy; CIU/min = Correct Information Units/min; Picture naming = percent of correctly named pictures out of 60 (Boston Naming Test); Post40 refers to assessment after 40 treatment sessions; Post75 = after 75 treatment sessions.
(yellow level of activation) and the slightly lower magnitude of activation in the left posterior perisylvian region (more red than yellow) comparing the images after MIT with the images after SRT treatment for Patient #2.

The between-treatments comparison (Patient #1 MIT vs. Patient #2 SRT) made after 40 sessions (see also Table 1) showed that the MIT-treated patient had greater improvement on all outcomes than the SRT-treated patient. fMRI studies revealed that Patient #1 showed significant fMRI changes in a right-hemisphere network involving the premotor, inferior frontal, and temporal lobes, while Patient #2 had changes in a left-hemisphere network consisting of the inferior pre- and post-central gyrus and the superior temporal gyrus.

Following his post40-SRT assessment, Patient #2 was enrolled in the MIT treatment, and the post40 scores became the new baseline from which the effects of MIT would be measured. After 40 sessions of MIT, Patient #2 showed a further increase in speech output and picture-naming, and his post75-MIT assessments revealed further gains in speech output while the picture-naming score remained stable (see Table 1).

**Discussion**

The traditional explanation for the dissociation between speaking and singing in aphasic patients is the presence of two routes for word articulation: one for spoken words through the brain’s left hemisphere, and a separate route for sung words that uses either the right or both hemispheres. The small amount of empirical data available supports a bihemispheric role in the execution and sensorimotor control of vocal production for both speaking and singing (Bohland & Guenther, 2006; Brown, Martinez, Hodges, Fox, & Parsons, 2004; Guenther, Hampson, & Johnson, 1998; Jeffries, Fritz, & Braun, 2003; Ozdemir et al., 2006), with a tendency for greater left-lateralization for speaking under normal physiological conditions (i.e., faster rates of production during speaking than singing). The representation of sensory elements of music and language might be either separate, or in different locations with smaller degrees of overlap (for more details on this see also Koelsch, Fritz, Schulze, Alsop, & Schlaug, 2005; Koelsch, Gunter, von Cramon, Zysset, Lohmann, & Friederici, 2002; Patel, 2003; Peretz, 2003). Nevertheless, if there is a bihemispheric representation for speech production, then the question of why an intervention that uses singing or a form of singing such as MIT has the potential to facilitate syllable and word production, still remains. In theory, there are four possible mechanisms by which MIT’s facilitating effect may be achieved: (1) **Reduction of speed:** in singing, words can be articulated at a slower rate than in speaking, thereby reducing dependence on the left-hemisphere; (2) **Syllable lengthening:** provides the opportunity to distinguish the individual phonemes that together form words and phrases. Such connected segmentation, coupled with the reduction of speed in singing, can help nonfluent aphasic patients become more fluent, and may receive greater support from right-hemisphere structures; (3) **Syllable “chunking”:** prosodic features such as intonation, change in pitch, and syllabic stress may help patients group syllables into words and words into phrases, and this “chunking” (Chase & Simon, 1973; de Groot, 1965) may also enlist more right-hemisphere support; and 4) **Hand tapping:** it is likely that MIT engages a right-hemispheric, sensorimotor network through the tapping of the patient’s left hand as each syllable is sung (one tap/syllable, one syllable/s), which may in turn provide an impulse for verbal production in much the same way that a metronome has been shown to serve as a “pacemaker” in other motor activities (rhythmic anticipation, rhythmic entrainment; Thaut, Kenyon, Schauer, & McIntosh, 1999). In addition, there may be a set of shared neural correlates that control both hand movements and articulatory movements (Gentilucci, Benuzzi, Bertolani, Daprati, & Gangitano, 2000; Meister, Boroojerdi, Foltys, Sparing, Huber, & Topper, 2003; Tokimura, Tokimura, Oliviero, Asakura, & Rothwell, 1996; Uozumi, Tamagawa, Hashimoto, & Tsuji, 2004), and further, the sound produced by the tapping may encourage auditory-motor coupling (Lahav, Saltzman, & Schlaug, 2007).

The two unique elements of MIT most likely to make the strongest contribution to the therapy’s beneficial effects are the melodic intonation with its inherent sustained vocalization, and tapping with the left hand. How might melodic intonation influence recovery? Functional imaging tasks targeting the perception of musical components that require a more global than local processing strategy (e.g., melodic contour, musical phrasing, and/or meter) tend to elicit greater activity in right-hemispheric brain regions than in left-hemispheric regions. It has been shown that tasks that emphasize spectral information over temporal information have shown more right- than left-hemispheric activation (Zatorre & Belin, 2001). Similarly, patients with right-hemisphere lesions have greater difficulty with global processing (e.g., melody and contour processing) than those with left-hemisphere lesions (Peretz, 1990; Schuppert, Munte, Wieringa, & Altenmueller, 2000). It is most likely that the two unique elements of MIT, the melodic intonation with its inherent sustained vocalization and the rhythmic tapping of the left hand, make the strongest contribution to the therapy’s beneficial effect.
The effects of the left hand tapping should be considered in the same context. Once the right temporal lobe is specifically engaged by the melodic intonation and melodic contour, it is conceivable that the role of the left hand tapping could be the activation and priming of a right-hemispheric sensorimotor network for articulation. Since concurrent speech and hand use occurs in daily life, and gestures are frequently used during speech, hand movements, possibly in synchrony with articulatory movements, may have a facilitating effect on speech production, but the precise role of this facilitation is unknown. We hypothesize that tapping the left hand may engage a right-hemispheric sensorimotor network that coordinates not only hand movements but orofacial and articulatory movements as well. There is some evidence in the literature that such superordinate centers exist in the premotor cortex and share neural substrates for hand and orofacial movements (Meister et al., 2003; Tokimura et al., 1996; Uozumi et al., 2004). Furthermore, behavioral (Gentilucci et al., 2000), neurophysiological (Meister et al., 2003; Tokimura et al., 1996) and fMRI studies (Aziz-Zadeh, Wilson, Rizzolatti, & Iacoboni, 2006) have shown that motor and linguistic cortical representations of objects are closely linked, and that the premotor cortex may belong to an integrative network coordinating motor and linguistic expression. An additional or alternative explanation is that the left hand tapping may serve the same function as a pacemaker or metronome has in rehabilitation of other motor activities, and in so doing, may facilitate speech production through rhythmic anticipation, rhythmic entrainment, or auditory-motor coupling (see also Lahav et al., 2007, and Thaut et al., 1999).

In summary, the melodic intonation and left hand tapping are the critical, unique elements of MIT that may likely be responsible for its therapeutic effect and might explain the predominant right hemispheric activation pattern seen in our prototypical patient. Elements of MIT that are shared with other, non-intonation-based therapies (e.g., the intensity of the intervention, direct therapist/patient interaction, unison and antiphonal repetition of words and phrases) also have therapeutic effects, as can be seen in our patient treated with SRT. Although caution should be exercised when making generalizations from results in two prototypical patients, we hope that our findings will serve as a source for further discussion on the efficacy of MIT, the neural correlates of MIT, and the choice of appropriate control interventions for MIT.

Author Note

This work was supported in part by grants from the National Institute of Neurological Disease and Stroke (NS045049, DC008796), the Doris Duke Charitable Foundation, the Grammy Foundation, and The Mattina R. Proctor Foundation.

Correspondence concerning this article should be addressed to Gottfried Schlaug, Department of Neurology, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215. E-MAIL: gschlaug@bidmc.harvard.edu

References


right-hemisphere facilitation in nonfluent aphasia.
Aphasiology, 15, 131-149.


GOTTFRIED SCHLAUG, SARAH MARCHINA, AND ANDREA NORTON, FIGURE 3.

fMRI activation maps (superimposed onto the surface projections of a spatially standardized normal brain) of the contrast “Overt Speaking vs. Silence (control condition)” (p < 0.05 FWE) for Patients #1 and #2 at baseline and after 40 therapy sessions. The color coding reflects different magnitudes of activation: the color yellow reflects a stronger activation than the color red. Patient #2 also shows the “Overt Speaking vs. Silence (control condition)” contrast after an additional 40 sessions of MIT that followed the SRT sessions. Slight differences can be seen in the regional magnitude of activation with a greater magnitude of activation in the right premotor/motor and temporal lobes (yellow level of activation) and the slightly lower magnitude of activation in the left posterior perisylvian region (more red than yellow) comparing the images after MIT with the images after SRT treatment for Patient #2.