Appendix A: Copies of treatment research studies included in Table 2

European Journal of Neuroscience, Vol. 26, pp. 2369–2375, 2007

doi:10.1111/j.1460-9568.2007.05810.x

Rhythmic auditory stimulation modulates gait variability in Parkinson's disease

Jeffrey M. Hausdorff,[1,2,3] Justine Lowenthal,[1,2] Talia Herman,[1] Leor Gruendlinger,[1] Chava Peretz[1,2] and Nir Giladi[1,2,4]

[1]Laboratory for Gait and Neurodynamics, Movement Disorders Unit and Parkinson Center, Department of Neurology, Tel Aviv Sourasky Medical Center, 6 Weizman Street, Tel Aviv 64239, Israel

[2]Department of Physical Therapy, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[3]Division on Aging, Harvard Medical School, Boston, MA, USA

[4]Department of Neurology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

Keywords: gait variability, Parkinson's disease, plasticity, rhythmic auditory stimulation

Abstract

Patients with Parkinson's disease (PD) walk with a shortened stride length and high stride-to-stride variability, a measure associated with fall risk. Rhythmic auditory stimulation (RAS) improves stride length but the effects on stride-to-stride variability, a marker of fall risk, are unknown. The effects of RAS on stride time variability, swing time variability and spatial-temporal measures were examined during 100-m walks with the RAS beat set to 100 and 110% of each subject's usual cadence in 29 patients with idiopathic PD and 26 healthy age-matched controls. Carryover effects were also evaluated. During usual walking, variability was significantly higher (worse) in the patients with PD compared with the controls ($P < 0.01$). For the patients with PD, RAS at 100% improved gait speed, stride length and swing time ($P < 0.02$) but did not significantly affect variability. With RAS at 110%, reductions in variability were also observed ($P < 0.03$) and these effects persisted 2 and 15 min later. In the control subjects, the positive effects of RAS were not observed. For example, RAS increased stride time variability at 100 and 110%. These results demonstrate that RAS enables more automatic movement and reduces stride-to-stride variability in patients with PD. Further, these improvements are not simply a by-product of changes in speed or stride length. After walking with RAS, there also appears to be a carryover effect that supports the possibility of motor plasticity in the networks controlling rhythmicity in PD and the potential for using RAS as an intervention to improve mobility and reduce fall risk.

Introduction

Previous studies have shown that rhythmic auditory stimulation (RAS) can improve the spatiotemporal features of gait in patients with Parkinson's disease (PD) (Rubenstein *et al.*, 2002; Lim *et al.*, 2005). When using RAS, administered in the form of a metronome, gait speed and stride length improved in PD patients in both 'on' and 'off' states (McIntosh *et al.*, 1997). When RAS was administered at a rate either equal to the patient's baseline step rate or 10% higher, it reduced the double support time and increased stride length (Freedland *et al.*, 2002). Similarly, gait speed increased when RAS was set higher than the usual step rate (Howe *et al.*, 2003; Willems *et al.*, 2006). The effects of RAS may also carry over to no-RAS walking; an increase in gait speed and stride length were observed after administration of RAS to PD patients for 3 weeks (Thaut *et al.*, 1996).

Although there is evidence indicating that RAS improves stride length and other spatiotemporal features of gait in PD, the effects of RAS on stride-to-stride variability and the consistency or rhythmicity of the gait pattern are largely unknown. *A priori*, one could argue that if key features of gait in PD, such as stride length and gait speed, improve then rhythmicity would also improve with RAS. Further, one could suggest that because RAS sets the pace, it can act like an external rhythm generator and help restore rhythmicity, thereby reducing stride-to-stride variability. Conversely, in PD, gait rhythmicity often behaves differently from stride length, gait speed and other parkinsonian features (Blin *et al.*, 1991; Hausdorff *et al.*, 1998; Schaafsma *et al.*, 2003; Frenkel-Toledo *et al.*, 2005a; Hausdorff, 2005). Thus, one could suggest that damage to the circuits that regulate rhythmicity may prevent these patients from walking with a consistent pattern, even in the presence of RAS.

Given the association between stride-to-stride variability and falls (Nakamura *et al.*, 1996; Hausdorff *et al.*, 2001; Schaafsma *et al.*, 2003; Hausdorff, 2005) and the importance of learning more about the motor control of PD and the potential clinical utility of RAS, the primary objective of the present study was to test the hypothesis that RAS reduces stride-to-stride variability in patients with PD. Following the example of previous investigations, which demonstrated that improvements in gait speed and stride length may be rate dependent (Freedland *et al.*, 2002; Howe *et al.*, 2003; Willems *et al.*, 2006), we studied the effects of RAS when it was set to 100 and 110% of the usual step rate. We also examined possible carryover effects, compared the effects in PD with those of a group of healthy controls and studied the effects of RAS on spatiotemporal parameters.

Materials and methods

Subjects

Twenty-nine patients with PD and 26 age- and sex-matched healthy controls volunteered to participate in the study. Potential participants were identified from the registry of the Movement Disorders Unit at

Correspondence: Dr Jeffrey M. Hausdorff, as above.
E-mail: jhausdor@bidmc.harvard.edu

Received 16 May 2007, revised 24 July 2007, accepted 4 August 2007

the Tel Aviv Sourasky Medical Center. Patients were recruited if they were on a stable antiparkinsonian medication regimen, were free of motor response fluctuations, had mild to moderate disease severity, i.e. Hoehn & Yahr stage II–III (Hoehn & Yahr, 1967) and were able to ambulate independently for at least 100 m. Subjects were excluded if they had other neurological or orthopedic conditions, if their hearing was impaired (as determined using a 124-cp tuning fork) or if they had dementia, as determined by a Mini-Mental State Exam score (Folstein *et al.*, 1975) less than 24. Subjects of similar age who were reported as being free of neurological, visual, vestibular and gait disturbances were recruited from the community to form the control group. The study was approved by the local human studies committee of the Tel Aviv Sourasky Medical Center and informed written consent was obtained from all participants before they entered the study, according to the Declaration of Helsinki.

Subject characteristics

General demographic data and fall history (number of falls in the previous year) were obtained. Cognitive function was assessed using the Mini-Mental State Exam and functional mobility was evaluated using the Timed Up and Go test (Podsiadlo & Richardson, 1991). Patients were evaluated with respect to the Hoehn & Yahr staging (Goetz *et al.*, 2004) (a measure of PD disease severity) and PD severity was evaluated using the Unified PD Rating Scale (Fahn *et al.*, 1987). The total Unified PD Rating Scale and its motor subscale (part III) were obtained.

Assessment of gait

After completing the assessment of subject characteristics, the effects of RAS on gait were examined under six conditions in the following order: (i) baseline (walking at a usual walking, comfortable pace without RAS); (ii) walking with RAS matched to the baseline cadence (RAS = 100% of the baseline step rate); (iii) walking at a comfortable pace without RAS (to examine any immediate carryover effect); (iv) walking with RAS at 110% of the baseline step rate; (v) walking at a comfortable pace without RAS (to examine any immediate carryover effect) and (vi) walking at a comfortable pace without RAS after a 15-min rest (to examine any delayed carryover effect). For each RAS condition, a metronome (Quark Metronome, Qwik Time^{TM}) was set to the desired step rate (based upon each subject's usual walking step rate). Subjects listened to the beat of the metronome and were told to try to match their walking (foot contact of each step) with the RAS. The walking distance for each condition was 100 m (i.e. four times along a 25-m level corridor); subjects were instructed to keep walking in time to the metronome when they turned at the end of the corridor. There was a 2-min break between each condition except between the fifth and sixth conditions, where there was a 15-min break to allow for evaluation of longer-term carryover effects. Before the first and second conditions, the participants underwent training in order to familiarize them with the walking environment and test equipment, and to allow them to practice walking with the RAS apparatus. Approximately 2 weeks later, 15 patients participated in a second session that followed the same protocol as detailed above but without the RAS intervention, in order to examine any possible within-session training effects.

A previously described computerized force-sensitive system was used to quantify gait rhythm, timing of the gait cycle (i.e. the stride time), swing time and stride-to-stride variability (Bazner *et al.*, 2000; Frenkel-Toledo *et al.*, 2005b; Yogev *et al.*, 2005). The system measures the forces underneath the foot as a function of time. It consists of a pair of shoes and a recording unit. Each shoe contains eight pressure-sensitive sensors that cover the surface of the sole and measure the vertical forces under the foot. The recording unit ($19 \times 14 \times 4.5$ cm; 1.5 kg) is carried on the waist. Plantar pressures under each foot are recorded at a rate of 100 Hz. Measurements are stored in a memory card during the walk and, after the walk, are transferred to a personal computer for further analysis. The following gait parameters were determined from the force record using previously described methods that filter out outliers, like those caused by turns, before further processing (Hausdorff *et al.*, 2001; Schaafsma *et al.*, 2003; Frenkel-Toledo *et al.*, 2005b): average stride time (and cadence), average swing time (as percentage of gait cycle), stride time variability and swing time variability. The measure of cadence (as determined from each subject's average stride time) during the first test was used to define RAS in subsequent tests. Variability measures were quantified using the coefficient of variation, e.g. stride time variability = 100 × (SD of stride time ⁄ average stride time). These measures were obtained for the left and right foot but, as they were highly correlated with one another, here we report values based only on the right foot (i.e. Pearson's correlation coefficients indicated moderate to strong correlations for all measures in all test conditions, based on the World Health Organization classification scheme). The average gait speed was determined by measuring the time that the subject took to walk the middle 8 m of the walkway and then averaging over all transversals. Each subject's average stride length was determined by multiplying the average gait speed by the average stride time. Although the portion of the walk in which average gait speed and average stride length were calculated was not identical to the steps in which the other measures were determined, all calculations were designed to reflect steady-state walking.

Statistical analysis

Descriptive statistics are reported as mean ± SD. We used the Student's *t* and chi-square tests to compare the PD and control subjects with respect to different background characteristics (e.g. age and gender). In order to estimate the effect of RAS, we applied mixed-effect models for repeated measures to evaluate within-group and between-group differences. For each gait variable, we applied a separate model where the dependent variable was the gait measure (a continuous one) and the independent variables were categorical, i.e. the group (PD patients or controls), walking condition (e.g. no RAS or RAS at 100%) and group × walking condition interaction term. The fixed factors in these models were group and walking condition, whereas the subject was the random factor. In each model, for the walking condition, the no-RAS baseline walk was considered as the reference category, inherent in the modeling procedure. Mixed-effect models were also applied in order to assess any possible training effect on each gait parameter (here the independent variable was the order of the walking condition). *P*-values reported are based on two-sided comparison. A *P*-value of 0.05 was considered statistically significant. All statistical analyses were performed using SAS 8.2 (Proc Mixed).

Results

Subject characteristics

The PD patients and the controls were similar with respect to age, gender, weight and height (Table 1). There was a significant difference between the groups for the Timed Up and Go test, fall history and Mini-Mental State Exam scores. Patients with PD took significantly

TABLE 1. Characteristics of the study participants: patients with Parkinson's disease and control subjects

	P-value	Controls ($n = 26$)	Patients ($n = 29$)
Age (years)	NS	64.6 ± 6.8	67.2 ± 9.1
Height (cm)	NS	168.69 ± 8.59	166.44 ± 7.64
Weight (kg)	NS	71.5 ± 11.1	70.3 ± 8.4
Gender (male% : female%)	NS	47 : 53	55 : 45
Timed Up and Go test (s)	< 0.001	9.3 ± 1.7	11.9 ± 3.4
Mini Mental State Exam	< 0.001	29.6 ± 0.8	28.3 ± 1.5
No. of falls in previous year	< 0.001	0.0 ± 0.0	1.2 ± 2.1

Data are presented as mean ± SD. For fall history, significant group differences are found by using a one-sample Kolmogorov-Smirnov test, a one sample t-test (i.e. comparing with 0.0) or transforming the data and comparing fall status (yes/no) between the two groups using chi-square analysis or Fischer's exact test ($P < 0.002$). NS, non significant.

longer to perform the Timed Up and Go test, reported more falls in the previous year and performed slightly worse on the Mini-Mental State Exam. In the patient group, the mean Hoehn & Yahr stage was 2.4 ± 0.4 and Unified PD Rating Scale total and motor scores were 24.9 ± 8.5 and 15.8 ± 4.5, respectively, indicating mild to moderate disease severity.

As shown in Table 2, under usual walking conditions (without any RAS), the average stride time was similar in the two subject groups. Gait speed, stride length and swing time were, however, significantly reduced in the subjects with PD ($P < 0.001$). Without RAS, subjects with PD also walked with greater stride-to-stride variability, as reflected by significantly higher (i.e. worse) stride time variability and swing time variability compared with the control group.

Immediate effects of rhythmic auditory stimulation

Table 2 summarizes the effects of RAS on the gait of the two subject groups. When RAS was set to each subject's most comfortable walking step rate (i.e. with RAS equal to 100% of the non-RAS step rate, 111 steps/min, on average), mean stride times were unchanged relative to the no-RAS condition in both groups, as expected. For the subjects with PD, gait speed, stride length and swing time significantly increased with RAS. There were, however, no significant effects on the two measures of variability. For the control subjects, gait speed, stride length and swing time were unchanged with RAS at 100%. Surprisingly, stride time variability of the control subjects increased (became worse) compared with the no-RAS condition.

When RAS was set to 110% of the usual walking step rate (e.g. 122 instead of 111 steps/min), the average stride time was significantly reduced in both subject groups (see Table 2). For both groups, at this faster RAS, gait speed significantly increased compared with the no-RAS condition ($P < 0.001$). Stride length and swing time significantly increased in the patients with PD ($P = 0.02$) but were unchanged in the controls. Compared with the no-RAS condition, the two variability measures were significantly reduced (improved) in the patients with PD when walking with RAS at 110% of the most comfortable stepping rate. An example of the effect of RAS on stride time variability is shown in Fig. 1 for one patient with PD. In contrast, in the control subjects at 110% RAS, stride time variability was increased (worse) compared with the no-RAS condition, whereas swing time variability was unchanged.

Carryover effects

Table 3 summarizes the carryover effects of RAS. After walking with RAS set to 100% of the step rate, there was an immediate carryover effect for three parameters in the PD group; the increase in gait speed, stride length and swing time persisted even when RAS was removed. Consistent with the absence of any effect observed while walking with RAS set to 100% of the step rate, there was, however, no significant change in either measure of variability shortly after RAS was switched off. For the control group, carryover effects after walking with RAS set to 100% of their step rate were also similar to the effects observed when walking with RAS set to this step rate. There were no significant effects on stride time, swing time, stride length or gait speed but there was a small but significant increase in stride time variability compared with the usual walking, no-RAS condition.

After walking with RAS set to 110% of the step rate, carryover effects were also observed, both immediately and 15 min later (see Table 3). For the subjects with PD, all of the improvements observed while walking with RAS set to 110% were also observed 15 min later. This included a

TABLE 2. Effects of rhythmic auditory stimulation (RAS) at 100% and 110% of each subject's usual walking cadence (No RAS)

	Control subjects			Patients with PD		
	No RAS	RAS at 100%	RAS at 110%	No RAS	RAS at 100%	RAS at 110%
Stride time (s)	1.08 ± 0.09	1.09 ± 0.10 (NS)	1.01 ± 0.09 (0.001)	1.08 ± 0.13 (NS)	1.08 ± 0.13 (NS)	1.03 ± 0.15 (0.001)
Gait speed (m/s)	1.24 ± 0.14	1.24 ± 0.17 (NS)	1.33 ± 0.21 (0.001)	1.00 ± 0.21 (0.001)	1.04 ± 0.22 (0.02)	1.09 ± 0.25 (0.001)
Stride length (m)	1.35 ± 0.19	1.37 ± 0.23 (NS)	1.37 ± 0.25 (NS)	1.06 ± 0.21 (0.001)	1.10 ± 0.23 (0.02)	1.10 ± 0.25 (0.05)
Swing time (%)	36.3 ± 1.4	36.1 ± 1.8 (NS)	36.0 ± 1.9 (NS)	33.8 ± 3.3 (0.001)	34.3 ± 3.1 (0.006)	34.3 ± 3.2 (0.02)
Stride time variability (%)	1.8 ± 0.6	2.2 ± 0.8 (0.01)	2.2 ± 1.3 (0.03)	2.6 ± 1.0 (0.003)	2.4 ± 1.1 (NS)	2.2 ± 0.9 (0.004)
Swing time variability (%)	2.9 ± 1.3	2.8 ± 1.2 (NS)	3.1 ± 1.7 (NS)	4.7 ± 3.2 (0.01)	4.5 ± 3.4 (NS)	4.1 ± 3.1 (0.03)

Data are presented as mean ± SD; numbers in parentheses are P-values based on within-group comparisons to usual walking (i.e. without RAS), except for the usual walking for the Parkinson's disease (PD) group, where the P-values are with respect to the usual walking (no RAS) values of the control subjects. Note that, in response to RAS at 110% of usual walking cadence, significant effects in stride time variability were seen in both the patient and controls but the effects were in opposite directions. NS, non significant.

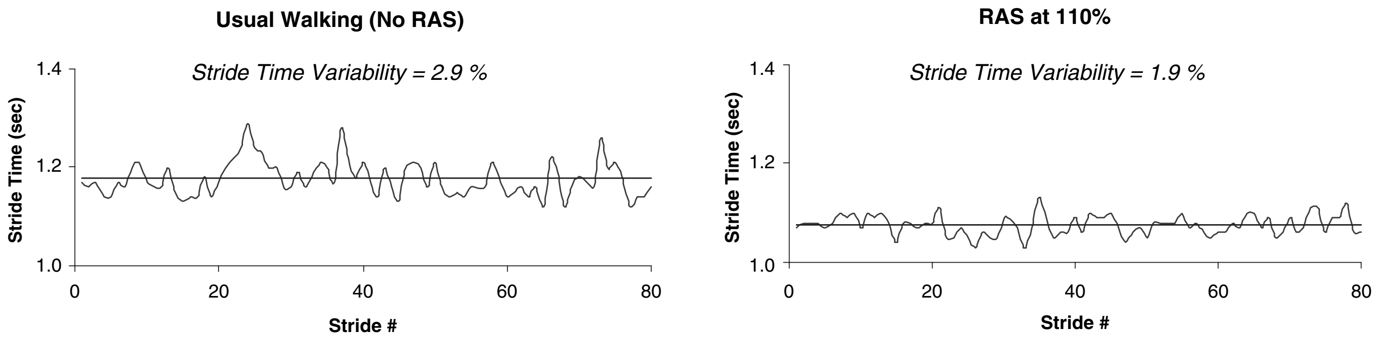


FIG. 1. Example of the effect of rhythmic auditory stimulation (RAS) (at 110% of the usual cadence) on the stride-to-stride variability of the stride time in one patient with Parkinson's disease. Note how the stride time fluctuates to a much lower degree with RAS (right, stride time variability 1.9%) compared with usual walking (left, stride time variability 2.9%). The horizontal lines are the average stride time for each condition.

TABLE 3. Carryover effects after walking with rhythmic auditory stimulation (RAS) set to 100 and 110% of each subject's usual walking cadence

	Control subjects			Patients with PD		
	15-min carryover after 110%	Immediate carryover effect at 110%	Immediate carryover effect at 100%	15-min carryover after 110%	Immediate carryover effect at 110%	Immediate carryover effect at 100%
Stride time (s)	1.05 ± 0.08 (0.003)	1.06 ± 0.09 (0.03)	1.07 ± 0.09 (NS)	1.03 ± 0.09 (0.001)	1.03 ± 0.09 (0.001)	1.06 ± 0.13 (NS)
Gait speed (m/s)	1.30 ± 0.16 (0.001)	1.29 ± 0.16 (NS)	1.27 ± 0.16 (NS)	1.11 ± 0.20 (0.001)	1.12 ± 0.19 (0.001)	1.06 ± 0.24 (0.001)
Stride length (m)	1.38 ± 0.20 (NS)	1.37 ± 0.20 (NS)	1.37 ± 0.20 (NS)	1.12 ± 0.20 (0.002)	1.14 ± 0.20 (0.001)	1.10 ± 0.23 (0.02)
Swing time (%)	36.1 ± 2.5 (NS)	36.0 ± 2.0 (NS)	35.9 ± 2.3 (NS)	34.9 ± 2.9 (0.001)	34.9 ± 2.5 (0.001)	34.4 ± 3.2 (0.009)
Stride time variability (%)	1.9 ± 0.5 (NS)	1.9 ± 0.5 (NS)	2.0 ± 0.7 (0.03)†	2.2 ± 1.0 (0.03)	2.4 ± 0.9 (NS)	2.5 ± 1.2 (NS)
Swing time variability (%)	2.9 ± 2.2 (NS)	2.6 ± 1.1 (NS)	2.8 ± 2.3 (NS)	3.5 ± 1.4 (0.003)	3.4 ± 1.8 (0.003)	4.2 ± 2.9 (NS)

Data are presented as mean ± SD; numbers in parentheses are *P*-values based on within-group comparisons to usual walking values (i.e. without RAS) shown in Table 2 (left-most column for each group). †Although RAS tended to reduce stride time variability in the subjects with Parkinson's disease (PD), it produced an increased stride time variability in the control subjects. NS, non significant.

significant reduction in both the stride time variability and swing time variability compared with pre-RAS walking. For the control subjects, the only changes that persisted 15 min after walking with RAS were the significant reduction in the average stride time and significant increase in the average gait speed. Significant carryover effects were not observed for either of the measures of variability in the control subjects.

As can be seen in Tables 2 and 3, the effects of RAS were similar in the two groups for certain gait measures but the effects were different in each group for others. Thus, for example, the effects of RAS on stride time and gait speed in the subjects with PD paralleled the effects observed in the control subjects ($P =$ NS for the group × intervention interactions). However, the effects on swing time and stride time variability were clearly different ($P \leq 0.01$ for the group × intervention interaction effects, e.g. RAS tended to reduce stride time variability in subjects with PD but to increase it in the control subjects).

Training effects

To test whether the observed effects of RAS may have been due to training, practice or familiarization with the surroundings, a subset of patients with PD ($n = 15$, arbitrarily selected) were asked to walk six 100-m bouts, similar to the RAS protocol but without any RAS. Compared with the first 100-m trial, there was a small but significant increase in gait speed and average stride time over time ($P < 0.0001$). In contrast, these repeated trials had no significant effects on either measure of variability (e.g. $P = 0.97$ for stride time variability).

Discussion

Consistent with previous reports (Blin *et al.*, 1990; Morris *et al.*, 1994; Hausdorff *et al.*, 1998; Ebersbach *et al.*, 1999a; Schaafsma *et al.*, 2003; Giladi *et al.*, 2005; Sofuwa *et al.*, 2005; Baltadjieva *et al.*, 2006), the patients with PD in this study walked with a reduced gait speed and swing time compared with an age-matched control group and with an increased stride-to-stride variability under the usual walking (no-RAS) condition. The effects of RAS on average gait speed, stride length and swing time were also generally similar to those reported earlier (McIntosh *et al.*, 1997; Freedland *et al.*, 2002; Rubenstein *et al.*, 2002; Howe *et al.*, 2003; Lim *et al.*, 2005; Willems *et al.*, 2006). For the patients with PD, RAS increased gait speed, stride length and swing time, both when RAS was set to the usual step rate and when it was 10% higher. The present investigation both concurs with and extends previous findings. The results are consistent with the reports that showed that RAS affects stride length and gait speed. Here we extend those studies by demonstrating that RAS also affects rhythmicity and stride-to-stride variability when it is set to a rate that is greater than the subject's usual cadence, that an intriguing carryover effect exists and that the effects of RAS on variability are apparently independent of those on stride length.

In the Introduction, we presented two possibilities about the effects of RAS on stride-to-stride variability; in brief, either it would reduce stride-to-stride variability or it would not. The results suggest that the answer is not a simple yes or no and that the stated arguments were not completely correct. Among the patients with PD, with RAS set to the usual walking step rate, both measures of stride-to-stride variability were not significantly different from those obtained while walking without RAS. Conversely, while walking at the higher step rate (110%), both measures of stride-to-stride variability were reduced (indicating improved rhythmicity and stability), moving closer to the control values. RAS seems to enhance rhythmicity in patients with PD but the effect may be rate dependent. In addition, the present findings demonstrate that the positive effects of RAS on variability are group specific and, interestingly, that the effects apparently endure even when RAS is removed.

How should the observed effects of the external cueing on stride-to-stride variability be interpreted? First, it is clear that, with appropriate cueing, patients with PD are capable of walking with a relatively reduced stride-to-stride variability and more 'normal' rhythmicity. Second, the differential effects of RAS on gait speed, stride length and variability in the subjects with PD (Table 2) suggest that the observed effects on stride-to-stride variability are not simply a by-product of changes in gait speed or stride length. At the lower RAS rate, gait speed and stride length of the patients with PD became larger than those of the no-RAS, usual walking condition. There was no effect on variability until RAS was increased to 110%, even though stride length remained unchanged compared with the value observed with RAS at 100%. Thus, the effects of RAS on variability are somewhat independent of the effects of RAS on stride length and gait speed.

Careful examination of the response of the control subjects and the effects of the repeated walks on gait speed but not on variability in PD also supports the idea that the RAS influence on variability differs from that of its impact on gait speed and stride length. Apparently, gait speed alone is not the single driver of variability, consistent with previous reports that observed a dissociation between stride length and variability (Frenkel-Toledo *et al.*, 2005a; Grabiner *et al.*, 2001; Hausdorff, 2004, 2005). Thus, the present results indicate that RAS apparently affects rhythmicity somewhat independently of speed and stride length. Although stride length and gait speed clearly play a critical role in many of the gait changes common in PD (Morris *et al.*, 1994, 1996), the present findings suggest that stride length does not fully determine rhythmicity, that RAS affects stride length and variability differently, and that the diminished stride length in PD is probably not the only source of the increased stride-to-stride variability in the presence of impaired basal ganglia function.

The effects of RAS on variability could be explained in several ways. Perhaps the most straightforward explanation for the observed effects of RAS on stride-to-stride variability is that, in patients with PD, RAS acts like a pacemaker and provides an external rhythm that is able to stabilize the defective internal rhythm of the basal ganglia (McIntosh *et al.*, 1997; Brotchie *et al.*, 1991; Thaut, 2003; Jantzen *et al.*, 2005; Zelaznik *et al.*, 2005; Nagy *et al.*, 2006). RAS may circumvent the pallidal- supplementary motor area pathway, possibly via the premotor cortex, and provide external cues to guide movement (Mushiake *et al.*, 1991; Halsband *et al.*, 1993; Hanakawa *et al.*, 1999a; Elsinger *et al.*, 2003). Increased activation of the lateral premotor cortex in PD patients during cueing lends support to this view (Hanakawa *et al.*, 1999b).

This explanation, that RAS works simply by acting as an external time-keeper, is rather intuitive but the rate-dependent nature of the effect of RAS on variability, the group-specific response and observed carryover effects present difficulties to this theory. If RAS acts as a pacemaker, it is not readily apparent why RAS increased stride-to-stride variability in the healthy control subjects, rendering their gait rhythm more 'abnormal' and why the effect should be rate dependent in PD. Indeed, further support for the rate-dependent nature of RAS comes from previous studies that showed that RAS at 80% of the normal cadence and similarly RAS at 60 beats/min increased stride time variability in patients with PD (Ebersbach *et al.*, 1999b; Almeida *et al.*, 2007). The carryover effects of RAS on variability (Table 3) pose the greatest challenge to the idea that RAS enhances rhythmicity simply by supplying an external cue. It is difficult to understand why the effects of RAS persist in the absence of the external cueing. Instead, the enduring effects suggest that neural time-keeping circuitry is apparently influenced by RAS.

Other explanations of the observed effects of RAS revolve around its influence on the neural circuitry regulating gait. The rate-dependent and enduring effects could be achieved by affecting striatal activation (Brown *et al.*, 2006) or via an effect on the cerebellum (Eckert *et al.*, 2005; Brown *et al.*, 2006; Del Olmo *et al.*, 2006). Indeed, extensive training of speech in PD has demonstrated evidence for striatal plasticity, changes in the cerebellum and long-term carryover effects that persist well beyond the training period (Liotti *et al.*, 2003), effects that may also be achievable with gait.

Just as extensive speech training has long-term benefits that have brought this approach into the clinic (Ramig *et al.*, 2001; Liotti *et al.*, 2003), the observed carryover effects raise the possibility that RAS might be used as a non-pharmacological intervention to complement standard pharmacological treatment in patients with PD. Indeed, RAS may have the potential to retrain some of the very basic and early abnormalities in PD, i.e. locomotion dysrhythmicity (Morris *et al.*, 1994; Ashburn *et al.*, 2001; Bloem *et al.*, 2004; Baltadjieva *et al.*, 2006), consistent with recent animal and human studies that support the idea that basal ganglia and related circuits manifest plasticity, even in the presence of PD (Liotti *et al.*, 2003; Fisher *et al.*, 2004; Wu & Hallett, 2005; Steiner *et al.*, 2006). The ability of RAS to restore rhythmicity and its previously demonstrated effect on gait speed and stride length (Thaut *et al.*, 1996; Howe *et al.*, 2003; Willems *et al.*, 2006) suggest that extensive use of RAS may have a profound, positive impact on mobility, fall risk and quality of life in PD (Morris *et al.*, 1994; Ashburn *et al.*, 2001; Bloem *et al.*, 2004). Important questions do, however, remain. The fact that improvement was observed even 15 min after walking with RAS was very encouraging and somewhat surprising. However, the carryover effects were acute and clinical significance remains to be proven. Nonetheless, the present findings combine with intriguing previous work (Thaut *et al.*, 1996; Howe *et al.*, 2003; Willems *et al.*, 2006) to argue for the plasticity of the neural control of gait rhythmicity in PD, independent of stride length, and for the use of RAS as an alternative therapeutic approach for the treatment of parkinsonian gait disturbances and, perhaps, for reducing fall risk.

Acknowledgements

We thank the subjects for their participation, time and effort, and Esther Eshkol for editorial assistance. This work was supported in part by NIH grants AG-14100, RR-13622, HD-39838 and AG-08812, by the US–Israel Binational Science Foundation, and by the National Parkinson Foundation, Miami, USA.

Abbreviations

PD, Parkinson's disease; RAS, rhythmic auditory stimulation.

References

Almeida, Q.J., Frank, J.S., Roy, E.A., Patla, A.E. & Jog, M.S. (2007) Dopaminergic modulation of timing control and variability in the gait of Parkinson's disease. *Mov. Disord.*, [Epub ahead of print, DOI: 10.1002/mds.21603.

Ashburn, A., Stack, E., Pickering, R.M. & Ward, C.D. (2001) A community-dwelling sample of people with Parkinson's disease: characteristics of fallers and non-fallers. *Age Ageing*, **30**, 47–52.

Baltadjieva, R., Giladi, N., Gruendlinger, L., Peretz, C. & Hausdorff, J.M. (2006) Marked alterations in the gait timing and rhythmicity of patients with de novo Parkinson's disease. *Eur. J. Neurosci.*, **24**, 1815–1820.

Bazner, H., Oster, M., Daffertshofer, M. & Hennerici, M. (2000) Assessment of gait in subcortical vascular encephalopathy by computerized analysis: a cross-sectional and longitudinal study. *J. Neurol.*, **247**, 841–849.

Blin, O., Ferrandez, A.M. & Serratrice, G. (1990) Quantitative analysis of gait in Parkinson patients: increased variability of stride length. *J. Neurol. Sci.*, **98**, 91–97.

Blin, O., Ferrandez, A.M., Pailhous, J. & Serratrice, G. (1991) Dopa-sensitive and dopa-resistant gait parameters in Parkinson's disease. *J. Neurol. Sci.*, **103**, 51–54.

Bloem, B.R., Hausdorff, J.M., Visser, J.E. & Giladi, N. (2004) Falls and freezing of gait in Parkinson's disease: a review of two interconnected, episodic phenomena. *Mov. Disord.*, **19**, 871–884.

Brotchie, P., Iansek, R. & Horne, M.K. (1991) Motor function of the monkey globus pallidus. 1. Neuronal discharge and parameters of movement. *Brain*, **114**, 1667–1683.

Brown, S., Martinez, M.J. & Parsons, L.M. (2006) The neural basis of human dance. *Cereb. Cortex*, **16**, 1157–1167.

Del Olmo, M.F., Arias, P., Furio, M.C., Pozo, M.A. & Cudeiro, J. (2006) Evaluation of the effect of training using auditory stimulation on rhythmic movement in Parkinsonian patients – a combined motor and [18F]-FDG PET study. *Parkinson. Relat. Disord.*, **12**, 155–164.

Ebersbach, G., Sojer, M., Valldeoriola, F., Wissel, J., Muller, J., Tolosa, E. & Poewe, W. (1999a) Comparative analysis of gait in Parkinson's disease, cerebellar ataxia and subcortical arteriosclerotic encephalopathy. *Brain*, **122**, 1349–1355.

Ebersbach, G., Heijmenberg, M., Kindermann, L., Trottenberg, T., Wissel, J. & Poewe, W. (1999b) Interference of rhythmic constraint on gait in healthy subjects and patients with early Parkinson's disease: evidence for impaired locomotor pattern generation in early Parkinson's disease. *Mov. Disord.*, **14**, 619–625.

Eckert, T., Barnes, A., Dhawan, V., Frucht, S., Gordon, M.F., Feigin, A.S. & Eidelberg, D. (2005) FDG PET in the differential diagnosis of parkinsonian disorders. *Neuroimage*, **26**, 912–921.

Elsinger, C.L., Rao, S.M., Zimbelman, J.L., Reynolds, N.C., Blindauer, K.A. & Hoffmann, R.G. (2003) Neural basis for impaired time reproduction in Parkinson's disease: an fMRI study. *J. Int. Neuropsychol. Soc.*, **9**, 1088–1098.

Fahn, S., Elton, R. & members of the UPDRS development committee (1987) Unified Parkinson's disease rating scale. In Fahn, S., Marsden, C.D., Calne, D. & Goldstein, M. (Eds), *Recent Developments in Parkinson's Disease*. Macmillan Health Care Information, Florham Park, NJ, pp. 153–163.

Fisher, B.E., Petzinger, G.M., Nixon, K., Hogg, E., Bremmer, S., Meshul, C.K. & Jakowec, M.W. (2004) Exercise-induced behavioral recovery and neuroplasticity in the 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine-lesioned mouse basal ganglia. *J. Neurosci. Res.*, **77**, 378–390.

Folstein, M.F., Folstein, S.E. & McHugh, P.R. (1975) 'Mini-mental state'. A practical method for grading the cognitive state of patients for the clinician. *J. Psychiatr. Res.*, **12**, 189–198.

Freedland, R.L., Festa, C., Sealy, M., McBean, A., Elghazaly, P., Capan, A., Brozycki, L., Nelson, A.J. & Rothman, J. (2002) The effects of pulsed auditory stimulation on various gait measurements in persons with Parkinson's Disease. *Neurorehabilitation*, **17**, 81–87.

Frenkel-Toledo, S., Giladi, N., Peretz, C., Herman, T., Gruendlinger, L. & Hausdorff, J.M. (2005a) Effect of gait speed on gait rhythmicity in Parkinson's disease: variability of stride time and swing time respond differently. *J. Neuroeng. Rehabil.*, **2**, 23.

Frenkel-Toledo, S., Giladi, N., Peretz, C., Herman, T., Gruendlinger, L. & Hausdorff, J.M. (2005b) Treadmill walking as an external pacemaker to improve gait rhythm and stability in Parkinson's disease. *Mov. Disord.*, **20**, 1109–1114.

Giladi, N., Hausdorff, J.M. & Balash, Y. (2005) Episodic and continuous gait disturbances in Parkinson's disease. In Galvez-Jimenez, N. (Ed.), *Scientific Basis for the Treatment of Parkinson's Disease*. Taylor & Francis, London, pp. 321–332.

Goetz, C.G., Poewe, W., Rascol, O., Sampaio, C., Stebbins, G.T., Counsell, C., Giladi, N., Holloway, R.G., Moore, C.G., Wenning, G.K., Yahr, M.D. & Seidl, L. (2004) Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: Status and recommendations The Movement Disorder Society Task Force on rating scales for Parkinson's disease. *Mov. Disord.*, **19**, 1020–1028.

Grabiner, P.C., Biswas, S.T. & Grabiner, M.D. (2001) Age-related changes in spatial and temporal gait variables. *Arch. Phys. Med. Rehabil.*, **82**, 31–35.

Halsband, U., Ito, N., Tanji, J. & Freund, H.J. (1993) The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain*, **116** (1), 243–266.

Hanakawa, T., Fukuyama, H., Katsumi, Y., Honda, M. & Shibasaki, H. (1999a) Enhanced lateral premotor activity during paradoxical gait in Parkinson's disease. *Ann. Neurol.*, **45**, 329–336.

Hanakawa, T., Katsumi, Y., Fukuyama, H., Honda, M., Hayashi, T., Kimura, J. & Shibasaki, H. (1999b) Mechanisms underlying gait disturbance in Parkinson's disease: a single photon emission computed tomography study. *Brain*, **122** (7), 1271–1282.

Hausdorff, J.M. (2004) Stride variability: beyond length and frequency. *Gait Posture*, **20**, 304.

Hausdorff, J.M. (2007) Gait dynamics, fractals and falls: Finding meaning in the stride-to-stride fluctuations of human walking. *Hum. Mov. Sci.*, **26**, 555–589.

Hausdorff, J.M., Cudkowicz, M.E., Firtion, R., Wei, J.Y. & Goldberger, A.L. (1998) Gait variability and basal ganglia disorders: stride-to-stride variations of gait cycle timing in Parkinson's disease and Huntington's disease. *Mov. Disord.*, **13**, 428–437.

Hausdorff, J.M., Rios, D. & Edelberg, H.K. (2001) Gait variability and fall risk in community-living older adults: a 1-year prospective study. *Arch. Phys. Med. Rehabil.*, **82**, 1050–1056.

Hoehn, M.M. & Yahr, M.D. (1967) Parkinsonism: onset, progression and mortality. *Neurology*, **17**, 427–442.

Howe, T.E., Lovgreen, B., Cody, F.W., Ashton, V.J. & Oldham, J.A. (2003) Auditory cues can modify the gait of persons with early-stage Parkinson's disease: a method for enhancing parkinsonian walking performance? *Clin. Rehabil.*, **17**, 363–367.

Jantzen, K.J., Steinberg, F.L. & Kelso, J.A. (2005) Functional MRI reveals the existence of modality and coordination-dependent timing networks. *Neuroimage*, **25**, 1031–1042.

Lim, I., van Wegen, E., de Goede, C., Deutekom, M., Nieuwboer, A., Willems, A., Jones, D., Rochester, L. & Kwakkel, G. (2005) Effects of external rhythmical cueing on gait in patients with Parkinson's disease: a systematic review. *Clin. Rehabil.*, **19**, 695–713.

Liotti, M., Ramig, L.O., Vogel, D., New, P., Cook, C.I., Ingham, R.J., Ingham, J.C. & Fox, P.T. (2003) Hypophonia in Parkinson's disease: neural correlates of voice treatment revealed by PET. *Neurology*, **60**, 432–440.

McIntosh, G.C., Brown, S.H., Rice, R.R. & Thaut, M.H. (1997) Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. *J. Neurol. Neurosurg. Psychiat.*, **62**, 22–26.

Morris, M.E., Iansek, R., Matyas, T.A. & Summers, J.J. (1994) The pathogenesis of gait hypokinesia in Parkinson's disease. *Brain*, **117** (5), 1169–1181.

Morris, M.E., Iansek, R., Matyas, T.A. & Summers, J.J. (1996) Stride length regulation in Parkinson's disease. Normalization strategies and underlying mechanisms. *Brain*, **119** (2), 551–568.

Mushiake, H., Inase, M. & Tanji, J. (1991) Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. *J. Neurophysiol.*, **66**, 705–718.

Nagy, A., Eordegh, G., Paroczy, Z., Markus, Z. & Benedek, G. (2006) Multisensory integration in the basal ganglia. *Eur. J. Neurosci.*, **24**, 917–924.

Nakamura, T., Meguro, K. & Sasaki, H. (1996) Relationship between falls and stride length variability in senile dementia of the Alzheimer type. *Gerontology*, **42**, 108–113.

Podsiadlo, D. & Richardson, S. (1991) The timed 'Up & Go': a test of basic functional mobility for frail elderly persons. *J. Am. Geriatr. Soc.*, **39**, 142–148.

Ramig, L.O., Sapir, S., Countryman, S., Pawlas, A.A., O'Brien, C., Hoehn, M. & Thompson, L.L. (2001) Intensive voice treatment (LSVT) for patients with Parkinson's disease: a 2 year follow up. *J. Neurol. Neurosurg. Psychiat.*, **71**, 493–498.

Rubenstein, T.C., Giladi, N. & Hausdorff, J.M. (2002) The power of cueing to circumvent dopamine deficits: a review of physical therapy treatment of gait disturbances in Parkinson's disease. *Mov. Disord.*, **17**, 1148–1160.

Schaafsma, J.D., Giladi, N., Balash, Y., Bartels, A.L., Gurevich, T. & Hausdorff, J.M. (2003) Gait dynamics in Parkinson's disease: relationship to Parkinsonian features, falls and response to levodopa. *J. Neurol. Sci.*, **212**, 47–53.

Sofuwa, O., Nieuwboer, A., Desloovere, K., Willems, A.M., Chavret, F. & Jonkers, I. (2005) Quantitative gait analysis in Parkinson's disease: comparison with a healthy control group. *Arch. Phys. Med. Rehabil.*, **86**, 1007–1013.

Steiner, B., Winter, C., Hosman, K., Siebert, E., Kempermann, G., Petrus, D.S. & Kupsch, A. (2006) Enriched environment induces cellular plasticity in the adult substantia nigra and improves motor behavior function in the 6-OHDA rat model of Parkinson's disease. *Exp. Neurol.*, **199**, 291–300.

Thaut, M.H. (2003) Neural basis of rhythmic timing networks in the human brain. *Ann. N.Y. Acad. Sci.*, **999**, 364–373.

Thaut, M.H., McIntosh, G.C., Rice, R.R., Miller, R.A., Rathbun, J. & Brault, J.M. (1996) Rhythmic auditory stimulation in gait training for Parkinson's disease patients. *Mov. Disord.*, **11**, 193–200.

Willems, A.M., Nieuwboer, A., Chavret, F., Desloovere, K., Dom, R., Rochester, L., Jones, D., Kwakkel, G. & Van Wegen, E. (2006) The use of rhythmic auditory cues to influence gait in patients with Parkinson's disease, the differential effect for freezers and non-freezers, an explorative study. *Disabil. Rehabil.*, **28**, 721–728.

Wu, T. & Hallett, M. (2005) A functional MRI study of automatic movements in patients with Parkinson's disease. *Brain*, **128**, 2250–2259.

Yogev, G., Giladi, N., Peretz, C., Springer, S., Simon, E.S. & Hausdorff, J.M. (2005) Dual tasking, gait rhythmicity, and Parkinson's disease: Which aspects of gait are attention demanding? *Eur. J. Neurosci.*, **22**, 1248–1256.

Zelaznik, H.N., Spencer, R.M., Ivry, R.B., Baria, A., Bloom, M., Dolansky, L., Justice, S., Patterson, K. & Whetter, E. (2005) Timing variability in circle drawing and tapping: probing the relationship between event and emergent timing. *J. Mot. Behav.*, **37**, 395–403.

ELSEVIER

Journal of Biomechanics 33 (2000) 1319–1323

JOURNAL OF BIOMECHANICS

www.elsevier.com/locate/jbiomech

Technical note

A measure of kinematic limb instability modulation by rhythmic auditory stimulation

Gary P. Kenyon, Michael H. Thaut*

Department of Electrical Engineering, Colorado State University, Center for Research in NeuroRehabilitation, Fort Collins, CO 80523, USA

Accepted 21 March 2000

Abstract

A mathematical method based on computations of residual absolute value sums (RAVS) was developed for the quantitative analysis of tremor-like perturbations of knee angle during the gait cycle. The method was tested on simulation data created by adding sinusoidal tremor of varying frequency and amplitude to the knee-angle graph of a healthy test subject. The method was then applied to compare knee tremor reduction, with and without auditory rhythm, in a group of five traumatically brain-injured patients with gait hemiparesis. Deviations from normal gait performance due to tremor were assessed by using self-comparison to a 17th-degree regression polynomial of each subject's own motion-, time-, and point-normalized knee- angle curve. With rhythmic cueing, the five subjects had a statistically significant RAVS-measured mean tremor reduction of $39.5 \pm 22.6\%$ ($t = -3.91$; $p = 0.017$). © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Tremor; Gait; Rhythm; Knee position analysis; Polynomial regression

1. Introduction

This study proposes a method of analysis for lower extremity tremor based on the measure of residual absolute value sums (RAVS) of the knee angle versus time curve. The method was used to quantify tremor and the tremor reduction resulting from rhythmic cueing of walking, based on self-comparison to the 17th-degree regression polynomial of the knee joint angle versus time curve. Auditory rhythm has been previously shown to improve gait kinematic parameters in subjects with hemiparetic gait disorders (Prassas et al., 1997; Thaut et al., 1997, 1999).

Applications of traditional waveform analyses to tremor data, e.g. accelerometry with Fourier analysis (Frost, 1978), have shown limited usefulness for a variety of reasons, including the difficulty of standardizing movement-related tremor for quantification (Deuschl et al., 1991), misinterpretation of multiple irregular peaks in a tremor spectrum (Gresty and Buckwell, 1990), low correlation with other tremor measures, the tendency to overestimate tremor improvement (Bain et al., 1993), the necessity of matching the data-processing algorithm to the data characteristics (Timmer et al., 1996), and the absence of information relating to the amplitude and duration of tremor or the coincidence of tremor with well-defined stages of a movement sequence. Furthermore, while Fourier analysis can yield useful information about the frequency content of the tremor, in gait the positional perturbation of the knee angle has functional importance for the stability of the movement. Therefore, a waveform analysis showing tremor reduction based on positional data would yield information clinically and functionally more useful for gait assessment and therapy.

2. Methods

2.1. Normalization of data

Trial variations in walking speed were normalized by expressing time duration as a fraction of one complete gait cycle. To normalize for variations in the knee-angle range of motion of a given subject, the value of the minimum joint angle over the gait cycle was subtracted from each data angle,

$$\theta_{\min} = \min\{\theta_j\}_{j=1}^{N}, \tag{1}$$

$$\tilde{\theta}_i = \theta_i - \theta_{\min}, \quad i = 1, 2, \ldots, N \tag{2}$$

* Corresponding author. Tel.: + 1-970-491-7384; fax: + 1-970-491-7541.

E-mail address: mthaut@lamar.colostate.edu (M.H. Thaut).

0021-9290/00/$ - see front matter
PII: S0021-9290(00)00077-4

and each shifted angle was then divided by the maximum shifted angle

$$\tilde{\theta}_{max} = \max\{\tilde{\theta}_j\}_{j=1}^{N}, \quad (3)$$

$$\hat{\theta}_i = \frac{\tilde{\theta}_i}{\tilde{\theta}_{max}}, \quad i = 1, 2, \ldots, N, \quad (4)$$

thus creating a dimensionless angle scale ranging in value from 0 to 1.

Each gait cycle was normalized to 1001 data points using an interpolation routine, since the RAVS procedure is dependent upon the number of points in the data sets. On the average, this increased the size of the data sets by a factor of four, and this level of interpolation has been shown to model human movement with finer resolution (Kenyon and Thaut, 2000).

2.2. Regression comparison reference

A tremor comparison reference was developed from a regressed version of each subject's personal knee-joint angle versus time curve, consistent with the assumption that tremor and other positional perturbations are superimposed on each subject's unique gait pattern. The choice of the degree for the regression polynomial was made on the basis of three procedures, each of which left the degree of the regression polynomial as a variable. First, the regression polynomial was tested for "goodness of fit" to a normal subject's knee-angle curve. The accuracy of the approximating regression polynomial was defined by Eq. (5), where ε represents the percent error measure, $\hat{\theta}$ is the normalized knee angle, and ρ is the corresponding knee angle after polynomial regression.

$$\varepsilon = \frac{\Sigma_{i=1}^{N}|\hat{\theta}_i - \rho_i|}{\Sigma_{i=1}^{N}|\hat{\theta}_i|} \times 100. \quad (5)$$

Second, it was required that the regression polynomial be a good approximation to the underlying non-perturbed gait patterns upon which tremor and other perturbations were superimposed. Third, the degree of the regression polynomial was "tuned" to give comparison results most consistent with the idealized changes within the tremor-simulated data sets, using tremor amplitude as the primary indicator of tremor severity.

2.3. Definition of the RAVS tremor measure M

The RAVS tremor measure M was defined as follows. First, a "residual" at a given data point was defined as the difference in value of the normalized knee angle and the regressed curve. The absolute values of the residuals were summed over the entire data set, then this sum was divided by the number of data points. In equation form

$$M = \frac{1}{N}\sum_{i=1}^{N} |\hat{\theta}_i - \rho_i| \quad (6)$$

where N = total number of normalized data points in the gait cycle. The resulting number may be thought of as the average tremor over the entire gait cycle. The percent change in tremor severity ΔT over two measurements (indicated by subscripts 1 and 2) was given by

$$\Delta T = \frac{M_2 - M_1}{M_1} \times 100. \quad (7)$$

2.4. Simulation

The validity of the RAVS measure M was tested by comparing changes in tremor severity, as calculated by Eq. (7), to tremor-amplitude changes in simulated data sets, which were constructed by mathematically adding sinusoidal tremor to the weight-bearing phase of one gait cycle of a 54-year-old male with normal gait. Ninety-seven simulations were created (including the zero tremor case) using tremor frequencies in the 4–15 Hz range (increments of 1 Hz), and tremor amplitudes in the 0.5–4.0° range (increments of 0.5°). The frequency range of 4–15 Hz includes the most widely documented results for tremor frequency, for a variety of tremor types, reported in the literature (Brooks et al., 1981; Deuschl et al., 1991; Elble et al., 1990; Frost, 1978; Ghika et al., 1993; Gresty and Buckwell, 1990; Jankovic and Frost, 1981; Marshall and Walsh, 1956; Timmer et al., 1996). The amplitude range, intended to represent barely noticeable to severe knee-joint tremors, was equal at its maximum to 6% of the normal subject's total knee-angle range, and included the tremor ranges of the five experimental subjects.

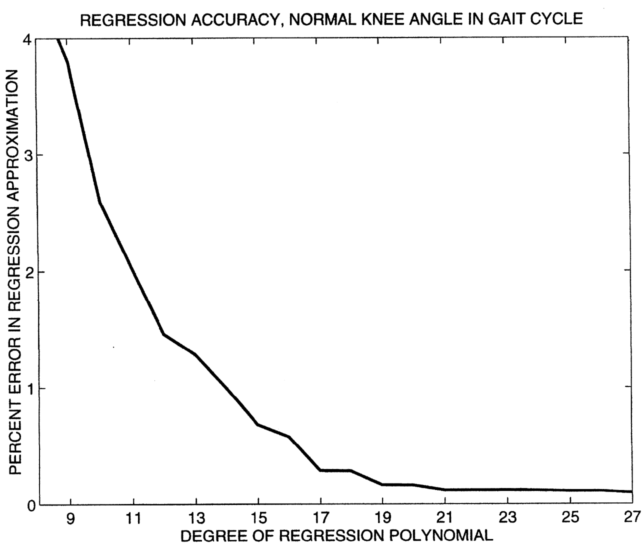


Fig. 1. This graph shows the error (as defined by Eq. (5)) in approximating the normal subject's knee angle over the gait cycle by a regression polynomial, with the degree of the regression polynomial as the independent variable. The approximation error is less than 1% when the degree of the regression polynomial is 15 or greater, and is monotonically decreasing.

Table 1
Summary of the results of applying the RAVS measure of tremor change to the 97 stimulated data sets (12 frequencies and 8 nonzero amplitudes, plus the zero amplitude case); (a) Shows tremor change calculated directly from amplitude change (Eq. (9)); (b) Shows mean tremor change over the entire frequency range as calculated by the RAVS method; (c) Shows the percent deviation of the Part (b) values from Part (a); (d) gives coefficients of variation for the percent deviations in Part (c).

A1	A2								
	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0
(a) Ideals (percent amplitude change)									
0.5	− 100.0	0.0	100.0	200.0	300.0	400.0	500.0	600.0	700.0
1.0	− 100.0	− 50.0	0.0	50.0	100.0	150.0	200.0	250.0	300.0
1.5	− 100.0	− 66.7	− 33.3	0.0	33.3	66.0	100.0	133.3	166.7
2.0	− 100.0	− 75.0	− 50.0	− 25.0	0.0	25.0	50.0	75.0	100.0
2.5	− 100.0	− 80.0	− 60.0	− 40.0	− 20.0	0.0	20.0	40.0	60.0
3.0	− 100.0	− 83.3	− 66.7	− 50.0	− 33.3	− 17.0	0.0	16.7	33.3
3.5	− 100.0	− 85.7	− 71.4	− 57.1	− 42.9	− 28.6	− 14.3	0.0	14.3
4.0	− 100.0	− 87.5	− 75.0	− 62.5	− 50.0	− 37.5	− 25.0	− 12.5	0.0
(b) Mean tremor change by RAVS method									
0.5	− 67.6	0.0	95.7	192.3	288.8	385.0	480.9	576.3	671.3
1.0	− 83.4	− 48.9	0.0	49.3	98.6	147.7	196.6	245.4	293.9
1.5	− 88.9	− 65.8	− 33.0	0.0	33.0	65.9	98.7	131.3	163.8
2.0	− 91.6	− 74.2	− 49.6	− 24.8	0.0	24.7	49.4	73.9	98.3
2.5	− 93.3	− 79.4	− 59.6	− 39.7	− 19.8	0.0	19.8	39.4	59.0
3.0	− 94.4	− 82.8	− 66.3	− 49.7	− 33.1	− 16.5	0.0	16.4	32.8
3.5	− 95.2	− 85.2	− 71.0	− 56.8	− 42.5	− 28.3	− 14.1	0.0	14.0
4.0	− 95.8	− 87.0	− 74.6	− 62.1	− 49.6	− 37.1	− 24.7	− 12.3	0.0
(c) Percent deviation of means from ideals									
0.5	32.4	0.0	− 4.3	− 3.9	− 3.7	− 3.8	− 3.8	− 3.9	− 4.1
1.0	16.6	2.2	0.0	− 1.4	− 1.4	− 1.5	− 1.7	− 1.9	− 2.0
1.5	11.1	1.4	0.9	0.0	− 0.9	− 1.6	− 1.3	− 1.5	− 1.7
2.0	8.4	1.0	0.7	0.8	0.0	− 1.1	− 1.2	− 1.5	− 1.7
2.5	6.7	0.8	0.6	0.7	0.8	0.0	− 1.2	− 1.4	− 1.7
3.0	5.6	0.7	0.6	0.7	0.7	3.0	0.0	− 1.8	− 1.6
3.5	4.8	0.6	0.5	0.6	0.9	1.1	1.4	0.0	− 1.8
4.0	4.2	0.6	0.5	0.7	0.9	1.1	1.3	1.5	0.0
(d) Coefficients of variation									
0.5	2.9	0.0	4.9	4.8	4.8	4.9	4.9	5.1	5.2
1.0	1.4	2.5	0.0	2.4	2.4	2.5	2.6	2.8	2.9
1.5	0.9	1.7	1.6	0.0	1.7	1.8	1.9	2.1	2.3
2.0	0.7	1.3	1.2	1.3	0.0	1.5	1.7	1.9	2.0
2.5	0.6	1.0	1.0	1.1	1.2	0.0	1.6	1.7	1.9
3.0	0.5	0.9	0.9	1.0	1.1	1.3	0.0	1.7	1.9
3.5	0.4	0.7	0.8	0.9	1.0	1.3	1.4	0.0	1.8
4.0	0.4	0.7	0.8	0.9	1.0	1.2	1.4	1.6	0.0

It was assumed that, under conditions of constant tremor duration, tremor amplitude is a valid indicator of the functional severity of tremor, and that a change in tremor severity is best reflected by a change in tremor amplitude. Thus, for the purposes of the simulated data sets, percent change in tremor severity corresponded to the percent change in tremor amplitude, and was calculated as

$$\Delta A = \frac{A_2 - A_1}{A_1} \times 100. \tag{8}$$

The results for the change in tremor ΔT calculated by the RAVS method were then compared to the idealized amplitude changes ΔA.

2.5. Pilot study with five subjects

The RAVS tremor measure was applied to data collected from five subjects to assess tremor activity and changes in tremor activity under the cueing conditions of with and without RAS. Subjects were 3 right-hemiparetic and 2 left-hemiparetic patients with traumatic brain injury (TBI) (3 female, 2 male; mean age 32 ± 7 yr). Two subjects had mild, 1 had moderate, and 2 subjects had severe lower limb spasticity (Brunnstrom, 1970). All subjects had visually observable tremors only during the weight-bearing phase of the stride cycle. RAS was presented free-field as a metronome click (1000 Hz, 20 ms plateau), which was frequency-matched to the step frequency recorded and computed for the trial without RAS. One full gait cycle was recorded (using 60-Hz videocameras) for three-dimensional digitized video motion analysis.

3. Results

3.1. Regression polynomial

The degree of the regression polynomial was set at 17 because at this value, the regression polynomial converged to the normal subject's knee-angle curve with less than 1% error (Fig. 1), as defined by Eq. (5). It also converged to the normal knee angle upon which tremor had been superimposed in the 97 simulation data sets, with a mean error of $0.70 \pm 0.65\%$. Degree 17 also resulted in the overall lowest values for the percent deviations from the simulation results (Table 1).

3.2. Subject group

For the subject group, the RAVS-measured tremor was reduced by $39.5 \pm 22.6\%$ (Table 2) when RAS was applied. This change was statistically significant using dependent sample *t*-tests ($t = -3.914$; $p = 0.017$) as well as a non-parametric analysis using the Wilcoxon Signed Ranks Test ($Z = -2.023$; $p = 0.043$).

3.3. Illustrative example

Graphs for one subject (Fig. 2) show the regression polynomial and the actual knee–angle curve for the cases of with and without RAS. Without RAS, the RAVS tremor measure has a value of 5.99 (by Eq. (6)), and with RAS a value of 2.98. Thus the resulting change in tremor, as calculated by Eq. (7), is -50.3%, indicating a 50.3% decrease in tremor.

Table 2
Individual and group results for the RAVS tremor measure computed for gait data collected without (M_1) and with (M_2) a Rhythmic Auditory Stimulus (RAS) for subjects with traumatic brain injury (TBI). Tremors ranging from very moderate ($M_1 = 0.76$) to severe ($M_1 = 5.99$) showed significant improvement with RAS

Subject	M_1	M_2	Tremor change (%)
1	0.76	0.26	− 66.1
2	5.12	4.80	− 6.3
3	0.80	0.56	− 30.1
4	2.55	1.41	− 44.7
5	5.99	2.98	− 50.3
Mean	3.04	2.00	− 39.5
Standard deviation	2.42	1.88	22.6

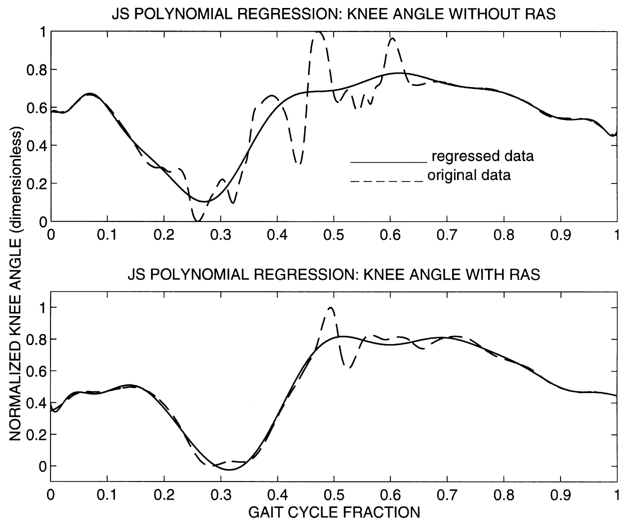


Fig. 2. For the two conditions of with and without rhythmic auditory stimulus (RAS), the dashed lines indicate the right knee angle over one complete gait cycle and the solid lines are the corresponding degree 17 regression polynomials. The RAVS method in essence measures the change, from one case to another, of the magnitude of the areas between the solid and dashed lines.

4. Discussion

The RAVS method applied to our study sample quantified and confirmed that tremor was reduced during rhythmic cueing. Considering motor control as an optimization problem of spatiotemporal precision (Harris and Wolpert, 1999), trajectory smoothing during RAS may be explained as a consequence of adding temporal stability and precision to the movement through the entrainment of the motor response to the rhythmic timekeeper.

The RAVS method, derived as a mathematical tool to quantify tremor severity, showed good results in computing tremor strength and changes in tremor as well as being sensitive enough to analyze degrees of tremor severity from very mild to very pronounced. Quantification of tremor change is an essential and new feature of this method, since no measure of tremor reduction has been found in the literature. Furthermore, unlike acceleration-based analyses, RAVS analysis uses position data itself rather than data derived from position, which allows analysis of duration and amplitude of tremor, and of correlations of tremor activity with movement landmarks of functional significance.

From a purely theoretical point of view, it is noteworthy that a direct mathematical correspondence exists between the amplitude of pure sinusoidal tremor and its RAVS-measure. For a tremor of frequency f Hz and amplitude A, with Δt as the time interval between RAVS data points, it can be shown, by relating both tremor amplitude and the RAVS measure of tremor to the area contained by the tremor curve, that the RAVS measure is proportional to the tremor amplitude:

$$M = \frac{2A}{N\pi f \Delta t}. \tag{9}$$

This result suggests that changes in tremor ΔT computed by the RAVS method should correspond closely with changes in tremor amplitude, assuming constant duration and location of tremor (i.e., weight-bearing phase) over trials.

The RAVS method may be particularly useful when analyzing pathological gait, because comparisons are made to each subject as their own "normalized" reference instead of comparing pathological gait patterns to averaged gait norms established with other study groups. Although the patient group of this study had tremor confined to the stance phase of the gait cycle, the RAVS method was also tested against simulation data sets with variable tremor duration over the entire gait cycle, with similar results as for the variable-amplitude simulations.

References

Bain, P.G., Mally, J., Gresty, M., Findley, L.J., 1993. Assessing the impact of essential tremor on upper limb function. Journal of Neurology 241, 54–61.

Brooks, V.B., Brookhart, J.M., Mountcastle, V.B., 1981. Motor control. In: Geiger, S.R. (Ed.), Handbook of Physiology, Vol. II. American Physiology Society, Bethesda, MD, pp. 325–343.

Brunnstrom, S., 1970. Movement Therapy in Hemiplegia. Harper & Row, New York, pp. 343–344.

Deuschl, G., Blumberg, H., Lucking, C.H., 1991. Tremor in reflex sympathetic dystrophy. Archives Neurology 48, 1247–1252.

Elble, R.J., Sinha, R., Higgins, C., 1990. Quantification of tremor with a digitizing tablet. Journal of Neuroscience Methods 32 (3), 193–198.

Frost, J.D., 1978. Triaxial vector accelerometry: a method for quantifying tremor and ataxia. IEEE Transactions on Biomedical Engineering BME-25, 17–27.

Ghika, J., Wiegner, A.W., Fang, J.J., Davies, L., Young, R.R., Growdon, J.H., 1993. Portable system for quantifying motor abnormalities in Parkinson's Disease. IEEE Transactions on Biomedical Engineering 40, 276–283.

Gresty, M., Buckwell, D., 1990. Spectral analysis of tremor: understanding the results. Journal of Neurology, Neurosurgery and Psychiatry 53, 976–981.

Harris, C.M., Wolpert, D.M., 1999. Signal-dependent noise determines motor planning. Nature 394, 780–784.

Jankovic, J., Frost, J.D., 1981. Quantitative assessment of parkinsonian and essential tremor: Clinical application of triaxial accelerometry. Neurology 31, 1235–1240.

Kenyon, G.P., Thaut, M.H., 2000. The discretization and interpolation of human movement and their effects on reported acceleration. Submitted for publication.

Marshall, J., Walsh, E.G., 1956. Physiological temor. Journal of Neurology, Neurosurgery and Psychiatry 19, 260–267.

Prassas, S.G., Thaut, M.H., McIntosh, G.C., Rice, R.R., 1997. Effect of auditory rhythmic cuing on gait kinematic parameters of stroke patients. Gait and Posture 6, 218–223.

Thaut, M.H., Kenyon, G.P., Schauer, M.L., McIntosh, G.C., 1999. The connection between rhythmicity in brain function: implications for therapy of movement disorders. IEEE Engineering, Medicine and Biology 18, 101–108.

Thaut, M.H., Rice, R.R., McIntosh, G.C., 1997. Rhythmic facilitation of gait training in hemiparetic stroke rehabilitation. Journal of Neurological Sciences 151, 207–212.

Timmer, J., Lauk, M., Deuschl, G., 1996. Quantitative analysis of tremor time series. Electroencephalography Clinical Neurophysiology 101, 461–468.

Journal of Music Therapy, XLIV (3), 2007, 198–216

Effect of Rhythmic Auditory Stimulation on Gait Performance in Children with Spastic Cerebral Palsy

Eunmi Emily Kwak, MME, MT-BC

Michigan State University

The purpose of this study was to use Rhythmic Auditory Stimulation (RAS) for children with spastic cerebral palsy (CP) in a clinical setting in order to determine its effectiveness in gait training for ambulation. RAS has been shown to improve gait performance in patients with significant gait deficits. All 25 participants (6 to 20 years old) had spastic CP and were ambulatory, but needed to stabilize and gain more coordinated movement. Participants were placed in three groups: the control group, the therapist-guided training (TGT) group, and the self-guided training (SGT) group. The TGT group showed a statistically significant difference in stride length, velocity, and symmetry. The analysis of the results in SGT group suggests that the self-guided training might not be as effective as therapist-guided depending on motivation level. The results of this study support three conclusions: (a) RAS does influence gait performance of people with CP; (b) individual characteristics, such as cognitive functioning, support of parents, and physical ability play an important role in designing a training application, the effectiveness of RAS, and expected benefits from the training; and (c) velocity and stride length can be improved by enhancing balance, trajectory, and kinematic stability without increasing cadence.

Rhythm is an essential element of music. Through simple observations of human behavior, people recognize that rhythmic music, even something as simple as toe tapping, influences human beings to synchronize movement. Even though the effect of rhythm on outward physical movement has long been evident, recent advances in technology and a culmination of medical research have revealed the role of auditory stimulation as an internalized timekeeper for rhythmic patterned movements (Thaut, 2005). The purpose of this research is to use rhythmic auditory stimulation (RAS) for children with spastic cerebral palsy (CP) to determine functional outcome effectiveness in gait training for ambulation. RAS uses music as an external time cue to regulate body movement.

CP is a collection of motor disorders resulting from brain damage before the age of two, when brain development is relatively complete (Taylor, 1993). In the general U.S. population, the prevalence of CP is 1.4 to 2.4 per 1,000 persons (Pellegrino, 1997). According to the United Cerebral Palsy Association, it is estimated that there are more than 500,000 Americans with CP (Taylor, 1993). Prevalence of CP is not influenced by factors such as culture, ethnic group, or socioeconomic status.

Significant features of participants with CP are impaired movement in ambulation that involves gait and upper body coordination, as well as problems with balance. A normal gait requires at least 30 major muscles working at exactly the right times and with exactly the right force to take two steps. Common problems in gait for those with spastic CP include ineffective gait patterns, such as short stride length; asymmetrical gait, slowness, impairment in coordination, and unnecessary body movement. Because of lack of muscle use, the degree of muscle weakness (atrophy) progresses throughout a child's development. Physical therapy is often recommended for children with CP to maintain and improve their physical functions. RAS has been found to be effective in an adjunctive role or as a sole method to increase the effectiveness of traditional physical therapy for ambulation in adult rehabilitation settings. CP patients encounter difficulties with coordination and muscle control similar to those experienced by rehabilitation patients, which suggests that RAS may be beneficial if used to enhance traditional physical therapy treatments.

RAS is defined as:

> A neurologic technique using the physiological effects of auditory rhythm on the motor system to improve the control of movement in rehabilitation and therapy. RAS is mostly used in gait therapy to aid in the recovery of functional, stable, and adaptive walking patterns in patients with significant gait deficits due to stroke, Parkinson's disease, traumatic brain injury, effects of aging, or other causes. (Thaut, 2005, p. 139)

The key element of RAS is the phenomenon of auditory entrainment, that is, the body's ability to synchronize its movements rhythmically. External auditory activity is mediated by internal un-conscious perceptual shaping at the subcortical level, and can arouse and raise the excitability of spinal motor neurons mediated by auditory-motor circuitry at the reticulospinal level (Pal'tsev & El'ner, 1967; Rossignol & Melvill Jones, 1976; Thaut, 1997a). The human body is a creative and resourceful organism and one of its most promising features, with regard to RAS, is that the various components of the brain are not connected by means of only one pathway; therefore, the brain does not completely stop working when one part is damaged or injured. When a part of the brain does not work properly, the brain finds a way to compensate for compromised functioning.

Damage to the motor cortex in a person with CP disturbs the normal process for the motor control system; in addition, the lower function motor control system could also be affected by brain damage. When the gait pattern of a person with CP is not rhythmic, it is likely that the internal timekeeper is not working. In these types of situations, RAS has been used to help regulate the motor control system by stimulating lower-level brain functions of the basal ganglia, cerebellum, brain stem, and spinal cord for patients with Parkinson's and other diseases. No conclusive evidence has been published, however, for using RAS in a clinical setting for children diagnosed with CP.

The RAS model is well defined by the Center for Biomedical Research in Music (CBRM) at Colorado State University. Research at Colorado State has demonstrated RAS effects in adult rehabilitation settings based on studies with Parkinson's disease, stroke, Huntington's disease, and traumatic brain injuries (Hurt, Rice, McIntosh, & Thaut, 1998; McIntosh, Brown, Rice, & Thaut, 1997; McIntosh, Thaut, & Rice, 1996; Thaut, 1997b). There is a need, however, to determine the functional outcomes of RAS in a range of populations with physical disabilities, including those diagnosed with CP, mental retardation, developmental delay, and physical deterioration due to aging. Though there are many applications for RAS, this study compared the effectiveness of RAS enhanced ambulation with traditional ambulation training in children with CP.

Treatments such as drug therapy, nerve injections, and orthopedic surgery are commonly used to treat spasticity to improve gait performance; yet, to date, there is no singular treatment that is most promising in reducing movement problems (Taylor, 1993). Unpredictability of results plagues those who attempt to treat individuals with spastic CP. The nature of CP and the complexity of brain function disturbance make treatment outcomes difficult to predict. Often, drug therapy causes drowsiness as a side effect, and injections of alcohol, phenol, or botulinum toxin are only effective for 6 months or less. Sometimes, corrective orthopedic surgery results in deformity of the antagonistic muscles (Miller & Bachrach, 2006).

RAS is a promising option within music therapy because the application of rhythm may organize an individual's gait and improve gait patterns. RAS training may be particularly beneficial because it has no negative side effects. In addition, it is cost-effective when compared to other treatments, and can be used in conjunction with other treatment modalities, or as an independent treatment because it is a noninvasive procedure.

A review of the literature revealed only one previous study with seven cerebral palsy patients in a home training setting. Results indicate improved velocity, cadence, stride length, and symmetry, as well as kinematic improvements of knee and hip ranges of motion and trajectories (Thaut, Hurt, Dragon, & McIntosh, 1998). The current study attempts to establish applications for using RAS in daily training with a music therapist, or as self-guided training, for establishing the validity of the research modality, and for adding to the theoretical basis of RAS. Investigation of the effectiveness of RAS for children with CP will help researchers learn appropriate applications for RAS and define the possibilities and limitations of RAS in clinical settings. The research questions are as follows: Does RAS enhance physical therapy for children with spastic cerebral palsy? Specifically, does it influence cadence, increase stride length, increase velocity, and improve gait symmetry? Is there a difference between the control, therapist-guided training group (TGT), and self-guided training group (SGT)?

Method

Participants

Thirty participants who have spastic CP ranged in age from 6–20 years old were enrolled in Chungju Hae-Hwa School for children with physical disabilities in Korea and participated in physical therapy for gait training at school. All were ambulatory but needed to stabilize gait and gain more coordinated movement. The control group ($n = 10$) received conventional gait training from a physical therapist while a music therapist observed. The therapist-guided training group (TGT, $n = 10$) received conventional gait training enhanced by RAS from both a physical therapist and a music therapist. The self-guided training group (SGT, $n = 10$) received conventional gait training from a physical therapist, and RAS self-guided training; a music therapist observed. Due to their illnesses, inconsistent school attendance, and lack of a posttest, five participants did not complete the experiment. Therefore, results were studied based on nine participants each in the control group and the TGT group, and seven participants in the SGT group.

Setting

For the pre and posttests, a 14-meter walkway was marked at 10 meters for test measurement with a 2-meter allowance for acceleration and 2-meter allowance for deceleration. Unrestricted space such as a playground, hallway, and slope (not stairs) at the school were used for gait training.

Equipment

Computer with cakewalk program. Cakewalk Pro Audio 8.0 is MIDI program which can provide variable tempo changes of recorded music used to accommodate the various cadences of each participant's gait. There were three basic melodies: "Dixie Land," "When the Saints Go Marching In," and a blues-style selection. All three songs have a basic steady beat pattern with 4/4 meter and have been recorded using quarter notes equal to 100 bpm. Since a majority of children with CP walk slower than typically developed children, the music was recorded slower than normal walking tempo (105 to 120 steps per minute). This slower tempo allows for a possible variation of tempo range from 80 bpm to 120 bpm.

Metronome. A metronome was used to confirm the accuracy of the tempo and assist in synchronizing participants during the warm-up activity in TGT and SGT.

Drums. A djembe was used by the therapist to emphasize the fundamental beat in the prescribed music. An open tone and a closed tone were used to produce different acoustics for the right foot and left foot. Clapping was also used for the same purpose. When a participant's cadence fell below 65 steps per minutes, to avoid excessively slow music, the cadence was multiplied by 2 as a tempo for the RAS music, and then the actual cadence was emphasized by drumming or clapping. For instance, a participant who walked 45 steps per minute had a music tempo of 90 beats per minute, and the drumming or clapping was maintained at 45 beats per minute.

Stride Analysis System. The accurate measurement of motor function and disability in patients is important in determining the efficacy of therapeutic interventions. Although clinical rating scales and simple timed tests of motor function are widely used to assess motor response to therapy, precise measurements for analyzing the effectiveness of music therapy interventions are necessary. The Stride Analyzer has been developed by B & L Engineering Inc. a group associated with Rancho Los Amigos Hospital (Norkin, 2000). The Stride Analyzer is a microprocessor/PC system designed to record foot-floor contact data by means of four pressure sensitive switches placed under the heel, at the heads of the first and fifth metatarsals, and at the big toe. It then calculates and compares gait parameters. Permanent records of the gait parameters and foot-floor contact patterns can be printed immediately following each test or later at the researcher's convenience. The Stride Analyzer is designed to analyze the following walking parameters: cadence, stride length, velocity, gait cycle, gait symmetry, and foot contact pattern.

Procedure

Pretest and posttest for all three groups. After all 30 participants signed consent forms, pretests were administered. During the pretest, while a participant was seated, the proper-sized footswitches were placed in the participant's shoes, and a recorder was attached to a belt placed around the participant's waist. Participants were asked to walk at their most comfortable tempo. When the participant walked into the test area, a manual trigger activated the data recorder and data was collected while the participant moved through the test area. Pretests were used as baseline data and for producing the prescribed music for each participant in the experimental groups. After completion of training, posttests were performed in the same way as pretests. All participants were asked to walk at their most comfortable tempo.

The prescribed music for RAS training. Based on the pretest, observation, and conference with the physical therapist, the tempo of the music was increased 5% above each participant's current walking tempo for the first week of training, or decreased, or retained at the current cadence depending on client's needs. The tempo was increased if balance and gait pattern tended to be better when a faster step was taken, just as riding a bicycle is easier at a faster speed. The tempo of the music was increased by 10 % from the baseline for the second week, and 15 % for the third week.

The increase in 5% from the current cadence was decided by a "Weber fraction". The "Weber fraction" is the percentage of the different thresholds obtained for different sensory stimulus. For example, in other to perceive the difference between electric shocks, a person needs to have 1.3% difference between them. In comparison, to taste sodium, 8.3% difference is needed. It was found that the Weber fraction for auditory time perception is 5% from 0.4 sec to 2.0 seconds. Beyond this range, the fraction is remarkably higher; however, the cadence of participants in this research was between 37 (1.62 seconds between steps) and 145 (0.41 seconds between steps) steps per minute. The imperceptible changes in the tempo of the music were essential to make training as comfortable as possible (Epstein, 1985; Getty, 1975; McBurney & Collings 1977). For the TGT group, the music was played through the computer's sound system. For the SGT group, the music was recorded on three cassette tapes and distributed to the participants. When a participant had a problem with gait posture such as toe walking due to muscle contracture, and/or spinal deformities, the tempo of the music was either decreased or maintained for the first week. Depending on their progress, tempo changes of the prescribed music were made every week.

Difference between the Therapist-Guided Training group (TGT) and the Self-Guided Training (SGT) Group. Both experimental groups used RAS technique for gait training, but the delivery of RAS technique was different; while the music therapist is actively involved with participants with verbal instructions and reinforcement, the music therapist was presented to SGT as an observer. Instead of warm-up exercise for the SGT group, muscle strengthening exercise using PSE and TIMP was took place. Other differences between TGT and SGT are illustrated in Table 1.

Procedure for RAS training for the Therapist-Guided Training Group (TGT). Sessions for each participant were held for 30 minutes five days a week, for three weeks. The general application for RAS training for TGT was consistent as possible, depending on a participant's ambulation ability, and total walking time. Individual organization for a session, total walking time and distance, and tempo

TABLE 1
Differences Between TGT and SGT

Difference	TGT	SGT
The presence of a music therapist	A music therapist is present and provides verbal instructions and reinforcement	A music therapist is present as an observer
The use of drum	A drum is used to emphasize the fundamental beat	None
Sound source	A computer speaker system is used to play the prescribed music during the entire session	Tape player and head phone set are used for playing the music
Environmental setting	One-to-one sessions in hallway or slope	Two–Three in a group and train together in hallway, slope or play ground
Muscle strengthening exercise	Depending on individual needs, physical therapist and the music therapist develop muscle strengthening exercises using PSE[a] and TIMP[b]	None

[a] Patterned Sensory Enhancement.
[b] Therapeutic Instrument Music Playing.

of the music were recorded and reported. During the entire session, prescribed music for the participant based on the Cakewalk program was played to increase the effect of entrainment. Basic organization for a 30-minute session was as follows.

Procedure for RAS training for the Self-Guided Training Group (SGT). At the first session, a tape and instructions were given to the participants and they were encouraged to participate during 3 weeks of training. The researcher demonstrated how they can feel the beat and how they can walk with the prescribed music. The music therapist was present during the self-training to record the duration of the training. The application below describes how sessions occurred. The therapist did not actively intervene with the training after the initial session. The participants were asked to follow the instructions and walk with the prescribed music for 30 minutes of daily self-training.

Procedure for the control group. A pretest and a posttest were administered for the participants in the control group. During the 3-week experiment, the participants in the control group had conventional physical therapy with the researcher present as an observer.

Results

During a 3-week research study, gait parameter, cadence, stride length, velocity, and symmetry ratio data were collected using the Stride Analyzer. In addition, pre and posttest data, and observations and information from parents and staff gathered during each treatment episode were recorded and used to answer the questions posed in this study.

Does RAS Training Influence Cadence?

Cadence was the element that the researcher used to adjust RAS training based on each participant's ability. Depending on the participants' gait performance, the treatment goals were to increase cadence in 9 cases, to decrease cadence in 4 cases and to maintain cadence in 3 cases. It was requested by the physical therapist at the school that some participants did not need to increase their cadence and some participant's cadence were faster than normal gait parameter due to their insecurity in balance, therefore unlike planned procedure, cadence needed to be adjusted depending on the participants. Cadence increased approximately 5% in the TGT and SGT groups and decreased by 1.2% in the control group (see Table 2). The results of a paired-sample *t*-test, however, indicate that there was no statistical difference between pre and posttest in the cadence within the groups (see Table 5).

Does RAS Training Increase Stride Length?

Stride length of a person with CP is often smaller than a normal person's stride length (Perry, 1992). Stride length improved approximately 15.8% overall. While the STG group showed only approximately 8% increase, the TGT group showed a 29.48% increase in stride length as shown in Table 3. The paired-samples *t*-test indicates that improvement of stride length in the TGT group was statistically significant ($t = -3.109$, $p = 0.014$), whereas the other groups did not show any significant differences (see Table 3).

Does RAS Training Increase Velocity?

Velocity, in general, improved from 20.73% (see Table 4) primarily in the TGT group. The results indicate improvement between pretests and posttests was 36.49% in the TGT group, 15.83% in the SGT group, and 9.44% in the control group. A paired-sam-

TABLE 2

Results of t *Test for Cadence Depending on Control, TGT and SGT Groups*

Group	*M*	*SD*	*t* value	2-tail significance
Control	1.1074	13.5551	0.245	0.813
TGT	–5.2667	28.0788	–.563	0.589
SGT	–5.2786	11.8615	–1.177	0.284

TABLE 3

Results of t *Test of Stride Length Depending on Control, TGT and SGT Groups*

Group	*M*	*SD*	*t* value	2-tail significance
Control	–5.2E-02	0.17948	–0.871	0.409
TGT	–0.20183	0.19474	–3.109	0.014*
SGT	–6.1E-0.2	0.19467	–0.829	0.439

* $p < 0.05$.

TABLE 4

Results of t *Test for Velocity Depending on Control, TGT and SGT Groups*

Group	*M*	*SD*	*t* value	2-tail significance
Control	– 2.8870	11.6944	–0.741	0.480
TGT	–11.1296	11.0245	–3.029	0.016*
SGT	–5.6179	9.2366	–1.609	0.159

* $p < 0.05$.

TABLE 5

Results of t *Test for Symmetry Depending on Control, TGT and SGT Groups*

Group	*M*	*SD*	*t* value	2-tail significance
Control	–8.0E-03	0.14251	–0.168	0.870
TGT	–0.13211	0.17020	–2.329	0.048
SGT	–8.0E-02	0.18258	–1.160	0.290

* $p < 0.05$.

ple t-test showed a difference between groups in improved velocity. The improved velocity in the TGT group was statistically significant ($t = -3.029$, $p = 0.016$), as illustrated in Table 4, whereas the other groups showed no significant difference. This indicated that the improvement in the TGT group was much greater than the improvement made by the other groups.

Does RAS Training Improve Symmetry?

Symmetry was calculated based on data collected by the Stride Analyzer. The shorter swing time of one leg from toe-off to heel strike was divided by the longer swing time of the other leg. Gait symmetry improved approximately 16.97% in the TGT group, and 9.92% in the SGT group, whereas the control group only increased by 0.91% as seen in Table 5. A paired-samples t-test indicated statistical significance in improving symmetry in the TGT group ($t = -3.029$, $p = 0.016$), whereas the other groups showed no significant difference. In addition to an increase in velocity and stride length, this indicates that the improvement in the TGT group is much greater than the improvement made by the other groups (see Table 5).

Is There A Difference between the Control, Therapist-Guided Training (TGT), and Self-Guided Training (SGT) Groups?

A one-way analysis of variance (ANOVA) was used to answer the research question regarding gait parameter improvement between the control, therapist-guided training (TGT), and self-guided training (SGT) groups. No significant difference between the groups was identified. This result is demonstrated in Table 6.

Data were analyzed using a paired-samples t-test, an independent-samples t-test, and an analysis of variance. In general, the therapist-guided training (TGT) group showed a significant difference

TABLE 6

Comparison Improvement of Gait Parameters between Control, TGT and SGT Groups

Variable	SS	df	MS	f	2-tail significance
Cadence	234.405	2	117.203	0.299	0.744
Stride length	0.123	2	6.141E-02	1.714	0.203
Velocity	315.476	2	157.738	1.346	0.281
Symmetry	7.847E-04	2	3.924E-04	1.292	0.295

in the paired-samples *t*-test within the therapist-guided training group. There were no significance difference on measures in other tests used for analysis; however, differences in velocity, cadence, and stride length were observable and indicated a positive outcome with the methods of this study (see Table 6).

Discussion

One of the objectives of the research was to define the details of RAS delivery techniques. During the 6-week intensive study, there were approximately 100 individuals with CP at Chungju Hae-Hwa School and only about 40 were ambulatory. The researcher observed their unique body movement; not only their gait but their whole-body movement patterns. It was obvious that each child had distinct and elaborate movement characteristics. Often, unique patterns were found in a child's body movement, as they used their left or right side. Individual gait problems were much more complex than the researcher expected. Influencing the treatment and results were deformation of the bone structure, abnormal muscle contracture at each of the major joints, balance, and equilibrium variations among individuals. General application for children with CP was difficult to establish, except the basic premise that the use of RAS music could stabilize and enhance their gait performance. Based on the result from the research, observation, and literature review, a schema for determining an individual RAS training application were developed (as shown in Figure 1).

Comments on the Research Procedure

During the experiment, it became clear that qualitative data collected could help clarify outcomes. The observation made by the researcher along with information from the school, medical personnel, and parents identified issues which impeded therapeutic gains. During the preparation for this study, concerns focused on participants' neurological damage and how RAS could affect gait training for persons with CP. The individual characteristics and personality of participants were not considered as factors when determining training application; however, it became clear during this study that certain characteristics in individual participants are essential in tailoring the training application and the expected outcomes. These characteristics included cognitive functioning, balance and equilibrium, assistive walking device, severity of condi-

A recommendation for RAS training application.

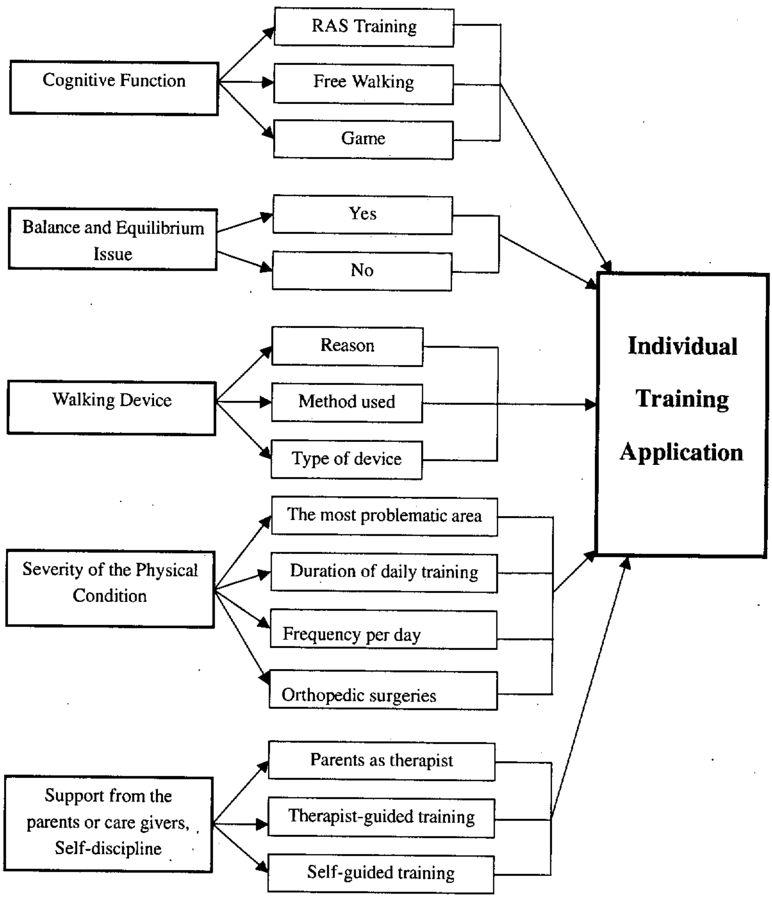


FIGURE 1.
A recommendation for RAS training application.

tion, history of orthopedic surgeries, personality, self-discipline, and support from parents or caregivers. Figure 1 illustrates some possible variables that influence a training application. The summary of the findings from this study follow.

A participant's cognitive function greatly influenced training application effects. The researcher found that Piaget's stages of cognitive theory readily applied to observations of the participants in the study (Smith, Cowie, & Blades, 2003). If the participant was cognitively under six-years old, the participant did not display strong self-discipline and the training was most successful when it was designed as an interesting game, such as pretending to be a soldier or walking to shop at imaginary stores (Berger, 1994). If a participant's cognitive function was less than one-year, free walking in a spacious room with a suitable sound presentation system was most appropriate for encouraging them to walk as they pursued a toy or an attractive object.

Balance and equilibrium issues required further investigation to improve participant's gait performance. If the participant suffered from these issues a set of exercises for enhancing balance and equilibrium reflex is required in the training application. They first need to learn to walk slowly by practicing "stepping and standing, stepping and standing" in order to improve a sense of equilibrium. Walking backward was especially helpful for balance and equilibrium. Exercise with inclining and declining slopes was useful to improve ability to balance and sustain equilibrium. In addition, a descending slope or stair is especially difficult for most people with a pathological gait. Practicing on a declining slope or stair encouraged the clients to cope with the fear of falling.

An assistive walking device necessitated an altered application of RAS. There are mainly two reasons for using assistive walking device: balance problems or muscle weakness problems. Consequently, the application was changed to accommodate their needs. In addition, there are different types of assistive walking devices, such as a four-wheel walker vs. two-wheel walker, and different methods for pushing a walking device such as lifting up a walker or rolling a walker. A person who has strong upper body strength excessively compensates their lower body weakness resulting in a deteriorating lower body condition. Those details were noted to tailor an individual application. It was important to incorporate an appropriate assistive walking device technique into gait training for participants to prevent further deterioration.

In consideration of all these factors, severity of the physical condition determines the types of muscle strengthening exercises, and duration of exercise time. Although the frequency of daily training during a day was kept to once a day because of their weaknesses, some participants showed the need to break down the daily training to twice a day. It is also important to accommodate a participant's orthopedic surgeries, especially if they have occurred within the past year. During the year, participants are in recovery; therefore rapid changes in body posture, recurrence of the spasticity of the operative muscles, and potential spasticity of the compensatory muscles are predicted. Depending on the changes, the training application was promptly adjusted.

Support from parents or caregivers, self-discipline, and personality of the participant seemed to influence outcomes, especially for the self-guided training group. Without strong support from the parents or caregivers and sheer determination of the participants, RAS training was less effective for the self-guided training group.

Recommendations

The most important factor in gait training for a person with CP is a comprehensive understanding of the person's pathological gait and scrupulous observation of their gait. To acquire information about the participant, contacts with the physical therapist, parents, and medical professionals are very beneficial. To understand their pathological gait patterns and pinpoint problematic muscles and patterns, it is imperative that the music therapist have a strong theoretical understanding of normal and pathological gait as well as clinical experience, and utilize basic measurement techniques.

Secondly, it is strongly recommended that when RAS training for a client with CP is designed and conducted, close cooperation should be established with the client's physical therapist because of the complexity of CP. To identify the effectiveness of RAS, the possible training application, and expected benefits from the training, individual characteristics such as cognitive functioning, support of parents, and physical abilities need to be taken into account. Participants' natural behavioral tendency (e.g., compliant, aggressive, negative) and participants' desire to improve also need to be taken into account for an individualized RAS training application. Third, duration and frequency of training must fit within a range of 10–20 minutes and up to twice a day based on the participant's physical condition. The fixed time of 30 minutes in this study was too long based on clients' laborious walking, especially for participants in the Self-Guided Training (SGT) group who solely walked as exercise for 30 minutes. A majority of participants requested shortening the duration of training, and even though the researcher encouraged them to finish the training tape, it often appeared to be laborious.

A final recommendation is to emphasize cadence. The relationship between increase in cadence and increase in velocity is important in application development for gait training. The results of RAS in Parkinson's disease, stroke, Huntington's disease, and traumatic brain injuries generally suggest that increased cadence would improve gait performance (Thaut et al., 1993; Thaut, McIntosh, & Rice, 1997; McIntosh et al., 1997; Miller & Bachrach, 1996). This does not appear applicable for spastic CP patients; in this study most participants enhanced their gait performance without increases in cadence. This may occur because a child with CP never learns to walk "correctly" or "normally." Unlike CP, the other populations studied experienced normal gait patterning prior to the onset of their illnesses and could have some undamaged cortical cells that could re-establish old motor pathways as well as form new motor pathways. Children with CP must rely on their damaged motor pathways, and RAS helps to develop new motor pathways. For the same reason, it is assumed that three weeks of RAS training for spastic CP patients is not sufficient to regulate their new gait pattern to become an ingrained gait pattern.

Future Research

It was found that using a drum with the prescribed music was very effective. However, because of clients' unstable gait, the therapist needed to be ready to assist to prevent falls or other needs. An assistant therapist or another helper should be available to play the drum. Clapping, however, worked very effectively in the current experiment to emphasize the downbeat, instead of a drum. When the clients' cadence were less than 65 steps per minutes, the prescribed RAS music was set at double the cadence to avoid excessively slow music, and drumming or clapping emphasized the actual cadence.

While the Stride Analyzer gives more details regarding foot contact pattern and allows researchers to understand overall body movements, some prospective participants might not be able to use it. Five prospective participants became uncooperative when the apparatus for the Stride Analyzer was attached to their shoes or feet were unable participated the experiment. Therefore, a manual measurement method needs to be considered to attain the gait parameters of participants through observations in relaxed environments.

During the experiment, the following three areas were identified as potential research projects. First, using patterned sensory enhancement (PSE) and therapeutic instrumental music playing (TIMP) were very useful in this experiment for strengthening and stretching muscles as a part of gait training but was not found to be documented in the literature. In this study, one of the major differences (beside therapist's present) between TGT and SGT was the muscle strengthening exercises. Initially, this component of gait training was added instead of warm-up exercise for SGT group, however, when the school physical therapist engaged with music therapy sessions, she really valued the potential possibility of using PSE and TIMP and used this part of RAS training to improve particular muscle groups. She was very impressed and pleased by the exercise progress and results. Therefore it is necessary to investigate the role of PSE and TIMP in RAS training.

Second, many individuals with CP cannot expect total independence. For instance, in special school of 101 students, only 32 walk independently, 24 use a wheelchair, and 45 use other walking devices. The neurologic effect of RAS for gait could transfer to upper body movement and posture reflex for CP through PSE and TIMP. The use of music combined with physical therapy for infants, toddlers, and adults with CP need to be examined.

Third, gait analysis in CP is usually focused on abnormal movement patterns, but in this study, four cases out of 25 had balance problems which were not considered in the research application. This, however, is an important factor to improve gait performance in children with CP. Many children with CP who are not able to walk often have difficulty sustaining normal posture reflexes that include the righting reactions and equilibrium reactions. This ability is essential to maintain balance and thus to achieve stable standing and walking. The possibility of the neurologic effect of auditory stimulation in problems with righting reactions and equilibrium reactions needs to be researched.

Conclusion

The purpose of this study was to use RAS with children with spastic CP in a clinical setting in order to determine its functional outcome effectiveness in gait training for ambulation. As mentioned above, a normal gait requires at least 30 major muscles working at exactly the right times and with exactly the right force to take two steps. Abnormalities in any one muscle can result in a pathological gait (Miller & Bachrach, 2006; Perry, 1992; Taylor, 1993). The effect of central nervous system lesions in mobility of a person with CP is unpredictable, and there are many variables that could affect the functional outcomes of different individuals' mobility, with the result that the gait pattern for each individual with CP is unique.

The exact mechanism and the neural basis of the synchronization of rhythmic physical movement such as finger-tapping or walking to an external auditory cue is not fully understood at present. The results of this study, however, support three conclusions:

RAS does influence gait performance of people with CP. Individual characteristics such as cognitive functioning, attitude, support of parents, and physical ability play an important role in designing a training application, the effectiveness of RAS, and expected benefits from the training, and stride length, velocity, and gait symmetry could be improved by enhancing balance, trajectory, and kinematic stability without increasing cadence.

Individuals have their own internal timekeeping mechanism for certain body movements. The analysis of the results from this study strongly suggests that increasing the cadence, in other words, changing the current internal timing, needs to be done very cautiously. Results suggest that it would be safer to maintain an individual's current cadence when an irregular foot contact pattern exists because this indicates balance and equilibrium problems, or when a pattern shows no floor contact on some part of the foot, which indicates muscle contracture or body deformation.

This research suggests that RAS may be very promising in gait training for children with CP. The matter at hand is to figure out how to apply it to their gait training. Continued research, however, is needed to explain how RAS enhances body posture and kinematic stability of gait performance.

References

Berger, K. S. (1994). *The developing person through the life Span* (3rd ed.). New York: Worth Publishers.

Epstein, D. (1985). Tempo relations: A cross-cultural study. *Music Theory Spectrum, 7,* 34–71.

Getty, D. (1975). Discrimination of short temporal intervals: a comparison of two models. Perception and Psychophysics 18, 1–8.

Hurt, C. P., Rice, R. R., McIntosh, G. C., & Thaut, M. H. (1998). Rhythmic auditory stimulation in gait training for patients with traumatic brain injury. *Journal of Music Therapy, 35,* 228–241.

McBurney, D. H., & Collings, V. B. (1977). *Introduction to sensation/perception.* Englewood Cliffs, NJ: Prentice-Hall.

McIntosh, G. C., Brown, S. H., Rice, R. R., & Thaut, M. H. (1997). Rhythmic auditory motor facilitation of gait patterns in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 62,* 22–26.

McIntosh, G. C., Thaut, M. H., & Rice, R. R. (1996). Rhythmic auditory stimulation as entrainment and therapy technique: Effects on gait of stroke and Parkinson's disease patients. In R. Pratt & R. Spintge (Eds.), *Music medicine: Vol. 2* (pp. 145–152). St. Louis, MO: MMB Music.

Miller, F., & Bachrach, S. J. (2006). *Cerebral palsy: A complete guide for caregiving* (2nd ed.). Baltimore, London: The Johns Hopkins Press Ltd.

Norkin, C. C. (2000). Gait analysis. In S. B. O'Sullivan & T. J Schmitz (Eds.), *Physical rehabilitation: Assessment and treatment* (4th ed.) (pp. 257–331). Philadelphia: F. A. Davis.

Pal'tsev, Y. I., & El'ner, A. M. (1967). Change in the functional state of the segmental apparatus of the spinal cord under the influence of sound stimuli and its role in voluntary movement. *Biophysics 12,* 1219–1226.

Pellegrino, L. (1997). Cerebral palsy. In M. L. Batshaw (Ed.), *Children with disabilities* (pp. 499–528). York, PA: The Maple Press Co.

Perry, J. (1992). *Gait analysis.* Thorofare, NJ: SLACK.

Rossignol, S., & Melvill Jones, G. (1976). Audiospinal influences in man studied by the H-reflex and its possible role in rhythmic movement synchronized to sound *Electroencephalography and Clinical Neurophysiology, 41,* 83–92.

Smith, P. K., Cowie, H., & Blades, M. (2003). *Understanding children's' development* (4th ed.), Oxford, Cambridge, Mass.: Basil Blackwell.

Taylor, F. (1993). Cerebral palsy: hope through research. In K. Bellenir (Ed.), (1997), *Congenital disorders sourcebook, Vol. 29* (pp. 119–177). Omnigraphics.

Thaut, M. H. (1997a). Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 63*(1), 22–26.

Thaut, M. H. (1997b). Rhythmic auditory stimulation in rehabilitation of movement disorders: A review of current research. In D. J. Schneck & J. K. Schneck (Eds.), *Music in human adaptation* (pp. 223–230). Blacksburg, VA: Virginia Polytechnic Institute and State University.

Thaut, M. H. (2005). *Rhythm, music, and the brain: Scientific foundations and clinical applications.* NY: Routledge.

Thaut, M. H., Hurt, C. P., Dragon, D. & McIntosh, G. C. (1998). Rhythmic entrainment of gait patterns in children with cerebral palsy [Abstract]. *Developmental Medicine and Child Neurology 40*(78), 15.

Thaut, M. H., McIntosh, G. C., & Rice, R. R. (1997). Rhythmic facilitation of gait training in hemiparetic stroke rehabilitation. *Journal of Neurological Science, 151,* 207–212.

Movement Disorders
Vol. 11, No. 2, 1996, pp. 193–200

Rhythmic Auditory Stimulation in Gait Training for Parkinson's Disease Patients

*M. H. Thaut, †G. C. McIntosh, †R. R. Rice, *R. A. Miller, *J. Rathbun, and *J. M. Brault

*Center for Research in NeuroRehabilitation, *Colorado State University and †Poudre Valley Hospital, Fort Collins, Colorado, U.S.A.*

Summary: Rhythmic auditory stimulation (RAS) was used as a pacemaker during a 3-week home-based gait-training program for Parkinson's disease (PD) patients (n = 15). Electromyogram (EMG) patterns and stride parameters were assessed before and after the test without RAS to evaluate changes in gait patterns. Data were compared with those of two control groups (n = 11), who either did not participate in any gait training or who participated in an internally self-paced training program. RAS consisted of audiotapes with metronome-pulse patterns embedded into the on/off beat structure of rhythmically accentuated instrumental music. Patients who trained with RAS significantly ($p < 0.05$) improved their gait velocity by 25%, stride length by 12%, and step cadence by 10% more than self-paced subjects who improved their velocity by 7% and no-training subjects whose velocity decreased by 7%. In the RAS-group, timing of EMG patterns changed significantly ($p < 0.05$) in the anterior tibialis and vastus lateralis muscles. Evidence for rhythmic entrainment of gait patterns was shown by the ability of the RAS group to reproduce the speed of the last training tape within a 2% margin of error without RAS. **Key Words:** Auditory rhythm—Gait training—Parkinsonism.

Gait deficits are among the most characteristic and most functionally debilitating signs of the motor neuropathology of Parkinson's disease (PD). Two of the most characteristic features of the gait profile of PD are bradykinesia and a shuffling stride pattern with shortened stride length and a reduced step cadence. Other changes include insufficient heelstrike and toe clearance, inadequate flexion about the hip, ankle, and knee, postural instability, and asymmetry between the stride times of the lower limbs (1).

Gait deficits in PD patients are often resistent to pharmacologic treatment despite the general effectiveness of dopaminergic drug therapy. In addition, prolonged drug intake may also be associated with decreased responsiveness to medication (2). Thus it is generally agreed that effective nonpharmacologic treatments need to be developed as an adjunct therapy to relieve symptoms and improve mobility (3).

Accepted September 12, 1995.
Address correspondence and reprint requests to Dr. M. H. Thaut at Department of Music and Department of Electrical Engineering, Colorado State University, Fort Collins, CO 80523, U.S.A.

In the clinical literature, the use of sensory systems through visual or auditory cues to facilitate locomotor activity has been repeatedly mentioned as one form of nonpharmacologic treatment (4). In fact, >25 years ago, Purdon Martin (5), in his classic treatise, described the use of visual cues to facilitate gait in PD. Despite some encouraging data, however, quantitative research into the controlled application of sensory cuing to motor facilitation with PD patients has been limited.

In previous work in sensorimotor facilitation, we were able to increase cadence, stride length, and symmetry in gait patterns of PD patients using rhythmic auditory stimulation (RAS) as peripheral timekeeper in a frequency entrainment design (6). These data were in agreement with findings by Richards et al. (7) for PD patients and by Thaut et al. for healthy individuals (8) and stroke patients (9). Decreased variability in rhythmic timing of arm and finger movements of PD patients using auditory rhythm was also reported by Freeman et al. (10) and Pastor et al. (11).

Considering (a) the evidence for beneficial effects of rhythmic cuing on motor performance, and (b) the need for effective alternative motor therapy strategies, we sought to determine in this study if possible entrainment effects of RAS on gait of PD patients could be used as an adjunct therapy strategy further to improve ambulatory function.

METHODOLOGY

Subjects

Study participants were volunteers who were randomly selected from referral lists from local PD support groups and primary care physicians. All subjects had a primary diagnosis of idiopathic PD. They all had significant gait deficits regarding velocity, stride length, and cadence (see description in Results) but were able to walk without physical assistance. The subject pool was randomly divided into an experimental group (EX) of 15 subjects (10 men and five women) and a control group, divided into a self (internally)-paced group (SPT; $n = 11$; eight men/three women) and a no-training group (NT; $n = 11$; eight men and three women).

Mean ages for the groups were EX, 69 ± 8 years; NT, 71 ± 8 years; SPT, 74 ± 3 years. Mean level of PD severity on the Hoehn and Yahr scale was 2.4 for EX, 2.6 for NT, and 2.5 for SPT. The EX group had a mean illness duration of 7.2 ± 4 years; the NT group, 8.5 ± 4 years; and the SPT group, 5.4 ± 3 years. The average number of reported falls per subject during the 12 months before the study was 4.5 for the EX group, 2.6 for the NT group, and 0.4 for the SPT group.

All patients were on a stable medication regimen (carbidopa/levodopa or selegiline/carbidopa and levodopa) during the study. No subjects were medicated with direct dopamine-receptor agonists such as bromocryptine or pergolide. Medication was monitored by a physician. All testing was done 90–120 min after first medication intake in the morning. No pronounced motor fluctuations or other cognitive, sensory, or mental health deficits were present in the study subjects.

Testing and Training

Subjects in each group participated in a pretest and posttest, 3 weeks apart, which involved measurement of their stride and electromyogram (EMG) patterns. Pretest and posttest performances were measured without the rhythmic timekeeper present. Subjects were instructed to walk at their normal speed.

During the training period, subjects in the NT group were instructed to just carry out their normal daily activities during the 3 weeks between tests. In the EX group, subjects exercised each day according to a prescribed program using RAS. The SPT group participated in the same exercise program but without the aid of RAS.

The EX subjects walked daily for 30 min with RAS. The RAS program consisted of walking on a flat surface, stair stepping, and stop-and-go exercises to rhythmically accentuated music at three different tempos. Subjects walked at each tempo for one third of the exercise time. Subjects could select from four short instrumental music pieces in four different styles familiar to the elderly age group in this study (folk, classical, jazz, country). Each selection was composed in 2/4 or 4/4 meter, and 32 measures in length. Rhythmic on-beats were enhanced by overlaying the click-function of the sequencer over the musical beat structure. We chose rhythmic stimuli embedded in a musical structure based on findings that rhythmic patterns within a musical context reduced response variability and synchronization offset more effectively than did single-pulse pattern in the frequency range of 1 to 2 Hz (60 to 120 steps/min) (12). Each selection was programmed on an eight-channel sequencer/synthesizer module in digital audio signal form (Alesys MMT8/Roland D5). Digital audiorecording allowed us to use the sequencer as a variable tempo driver to change the tempo of the music without loosing pitch control. The tempos were labeled "normal," "quick," and "fast."

For the first week of training, the normal tempo was the pretest cadence, the quick tempo was 5 to 10% faster, and the fast tempo an additional 5 to 10% faster. After each week, each tempo increased by 5 to 10%, so the quick tempo became normal, and so on. The rate of increase was based on the subjects' ability to match the tempos and the requirement to keep the fastest tempo from exceeding 130 steps/min. The subjects used portable tape players with light-weight headsets. Each musical selection was recorded for 30 min on tapes in the laboratory. The subjects exercised on their own or with spousal assistance at home or in the community.

The SPT group performed their walking sessions without RAS, following the same training protocol and training exercises for the same length of time. They were instructed to divide their exercise time into three equal periods, during which they would walk first at normal speed, then increase their speed for the next period, and increase it again for the last period. Both the EX group and the SPT group were visited once a week by a research assistant and also were asked to keep daily logs of their walking activities to control for training compliance. Compliance was 100% for both groups. One fall was reported by one EX-group subject during the study.

Data Recording

Gait was studied through measurement of footfall patterns and EMG recordings of the medial gastrocnemius (GA), tibialis anterior (TA), and vastus lateralis (VL) muscles on both sides, averaged across five strides. Data were recorded from subjects walking at their normal speed along a 6-m walkway with 2 m on each side for acceleration and deceleration. Two trials were run; the first was a flat surface walk followed by a walk over an incline-step obstacle. The incline-step obstacle was 3 m long with a 1-m incline rising 20 cm, a 1-m platform, a 10-cm step down, another 1-m platform, and a final 10-cm step down.

Data were recorded with an IBM-compatible PC and analog-to-digital converter, surface EMG electrodes, and a computerized foot-switch recording system. Foot-switch data recorded foot contact at the heel, first and fifth metatarsal, and big toe.

Raw EMG data were digitized at 500 samples per second. The digitized EMG was high-pass filtered at 70 Hz with a zero-phase finite-impulse-response filter and then full-wave rectified with a RMS (root-mean-square) processor. Smoothing of the RMS signal values was achieved by low-pass filtering with a 100-ms Hamming window impulse response, which had a −3-dB gain at 6.6 Hz. The average RMS value was then computed at 1% increments of the total gait cycle for every stride in the trial.

EMG Analysis

EMG data were analyzed through measurement of variability, symmetry, and timing of the muscle-activation period. Variability was measured as the weighted average of the coefficient of variability (COV) at each percentage of the gait cycle (GC; 13). Bilateral symmetry was measured as the correlation coefficient between amplitude-normalized average profiles (14). Timing of muscle activation was initially assessed by computing a timing power index (duty cycle index), which reflected the duration of muscle activation by using the power of the average waveform, normalized in amplitude and shifted to a zero-mean value. After this initial analysis, changes in the period of EMG activity were analyzed further through detection of onset and termination time by a computer algorithm.

Statistics

Analysis of variance (ANOVA) procedures with post hoc multiple comparisons were used to analyze statistical differences in change scores between pretest and postest in each group. Change score analysis was used to offset possible differences of study participants at pretest and to minimize group data variability. However, to assure that change score differences were not solely due to group differences at the outset of the experiment, separate pretest and posttest ANOVAs were also computed. To account for the small sample size, nonparametric ANOVA (Kruskal–Wallis) using rank-order transformed data was performed on all parameters in addition to parametric ANOVA.

RESULTS

Because parametric and nonparametric ANOVA procedures revealed the same results, only nonparametric data are reported here as the more stringent approach to identify statistical significance. Furthermore, ANOVA comparisons of all measures at pretest found no significant ($p < 0.05$) differences between groups.

Velocity, Cadence, and Stride Length

Velocity data were computed for the flat and the inclined walks. Cadence and stride length were computed only for the flat walk. The pretest and posttest results are listed in Table 1. During pretest, all subjects showed abnormal gait patterns characteristic of PD. These included decreased velocity (mean, 45.2 m/min), shortened stride length (mean, 0.98 m), and slow cadence (mean, 92.6 steps/min). The accepted normal age-matched values reported in the literature are 73 m/min for velocity, 1.27 m for stride length, and 113 steps/min for cadence (15). All experimental subjects increased their individual gait velocity during posttest (Fig. 1). The mean increase for the EX group was 24.1% (t = 3.84; p = 0.007) in the flat walk (48.7 m/min to 58.3 m/min) and 26.1% (t = 3.27; p = 0.009) over the incline-step obstacle (40.8 to 49.4 m/min). The velocity of the NT group actually decreased slightly from pretest to posttest (42.1 to 38.7 m/min), which suggests that familiarity with the laboratory environment was not a factor in the experimental group's improvement. The SPT group did show an improvement in velocity of 7.4% (47.9 to 51.9 m/min), which

TABLE 1. *Means, standard deviations, and percentage change score means for velocity, cadence, stride length*

	Velocity (m/min)			Cadence (steps/min)			Stride length (m)		
	Pre	Post	Change (%)	Pre	Post	Change (%)	Pre	Post	Change (%)
Flat walk									
EXP									
x̄	48.7	58.3	24.1	96.8	105.7	10.4	0.99	1.10	12.0
SD	13.6	12.6	12.3	13.5	11.5	13.5	0.19	0.17	12.8
NT									
x̄	42.1	38.7	−7.3	88.5	96.6	5.7	0.93	10.84	−10.3
SD	15.5	16	15.8	13.4	9.6	11.9	0.24	0.32	17.9
SPT									
x̄	47.9	51.9	7.4	92.6	92.2	−0.4	1.04	1.12	7.9
SD	7.3	12.5	10.7	7.6	8.4	1.7	0.16	0.21	9.2
Incline-step walk									
EXP									
x̄	40.8	49.4	26.1						
SD	14	13	30.9						
NT									
x̄	35.8	30.2	−10.5						
SD	15	13	20.9						
SPT									
x̄	40.7	43.9	8.4						
SD	9.3	11.6	16.2						

EXP group, n = 15; NT and SPT groups, n = 11.

is less than one third of the experimental group gain. The EX group's increase in velocity was nearly equally achieved through an increase in cadence and stride length, 10.4% (t = 2.9; p = 0.01) and 12.0% (t = 3.63; p = 0.009) respectively, whereas the SPT group's increase was achieved solely through increases in stride length of 7.9%. The SPT group's cadence remained unchanged from pretest to posttest. The pretest to posttest changes in both control groups were all statistically nonsignificant.

ANOVA comparisons between groups revealed significant differences in improvement of flat and incline velocity (FLAT: F = 13.77; p = 0.0001; INCLINE: F = 4.77; p = 0.0185). Post hoc comparisons showed that the EX group improved significantly over the SPT group (FLAT: p = 0.0307; INCLINE: p = 0.0347) and the NT group (FLAT: p = 0.0001; INCLINE: p = 0.0052). Changes in stride length and cadence were also significantly different between groups (STRIDE LENGTH: F = 4.86; p = 0.0144; CADENCE: F = 3.81; p =

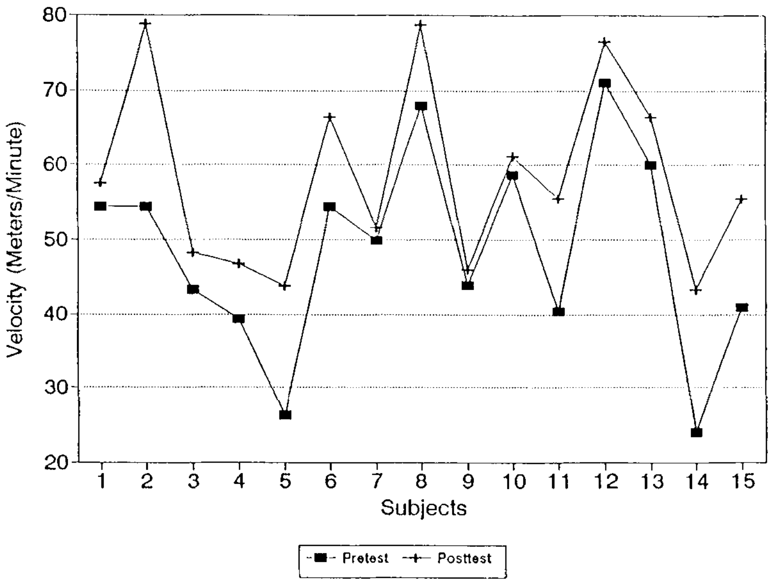


FIG. 1. Graph of gate velocity pretest (squares) and posttest (plus signs) for the 15 experimental subjects.

0.0328). However, in post hoc comparisons, the EX group's improvements in stride length were only significantly better than the NT group ($p = 0.0045$), and changes in cadence were only significantly higher than the SPT group ($p = 0.0340$).

To test for possible entrainment mechanisms between RAS-beat frequency and step frequency, subjects in the EX group were asked—after completion of the posttest—to reproduce the cadence of their fastest training tape from memory without RAS present. On average, the fastest training cadence was 18.2% faster than their normal posttest walk. The average absolute error in matching the training cadence was 5.0% (±2.9%), with the group average cadence only 1.8% (±5.6%) slower than the training tempo (Fig. 2).

The objective gains in gait performance were quite noticeable to the study subjects, who reported a 100% strong agreement on an exit questionnaire that RAS training had made their walking patterns more stable, had improved their speed, and had helped their walking in activities of daily life.

TABLE 2. *Means, standard deviations, and change score means for EMG variability (% amplitude ratio)*

	Gastrocnemius	Tibialis anterior	Vastus lateralis
Experimental group			
Pre			
$\bar{x}$	36.4	36.0	24.6
SD	12.0	10.0	5.1
Post			
$\bar{x}$	31.5	30.3	22.6
SD	9.8	10.4	7.3
Change			
$\bar{x}$	−4.9	−5.7	−2.1
SD	11.2	10.6	6.5
Control group (combined)			
Pre			
$\bar{x}$	35.0	43.1	28.7
SD	15.1	23.7	9.6
Post			
$\bar{x}$	31.6	36.6	27.2
SD	13.2	19.2	6.5
Change			
$\bar{x}$	−3.4	−6.5	−1.5
SD	8.4	10.5	5.9

EXP group, $n = 15$; CONTROL group, $n = 16$.

EMG Analysis

In each control group, three EMG records had to be rejected because of recording artifact. Thus data from the two groups were pooled to provide a sample size suitable for statistical analysis.

In Table 2, the percentage variability is listed for the GA, TA, and VL muscles. Pretest to posttest changes in the TA muscle approached statistical significance ($F = 2.29$; $p = 0.0559$), however, regardless of treatment condition. No changes in other muscles were found to be statistically significant.

EMG symmetry was examined by comparison of left and right side normalized average gait-cycle profiles in each muscle group, as listed in Table 3. In the EX group, the symmetry increased in each muscle pair. The posttest symmetry in each muscle was ~85%, which may be a practical limit due to variability in electrode placement (16). Symmetry in

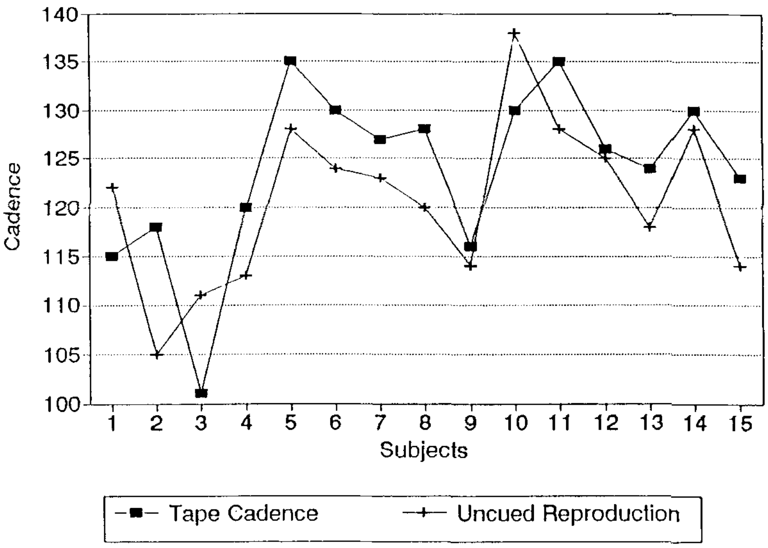


FIG. 2. Graph of maximal training tape cadence (squares) and uncued cadence reproduction (plus signs) for the 15 experimental subjects.

TABLE 3. *Means, standard deviations, and change score means for EMG symmetry (% symmetry ratio)*

	Gastrocnemius	Tibialis anterior	Vastus lateralis
Experimental group			
Pre			
x̄	73.0	71.5	85.1
SD	24.3	16.5	9.7
Post			
x̄	86.3	84.9	86.9
SD	8.2	7.7	9.0
Change			
x̄	13.3	13.4	1.8
SD	21.4	15.5	13.5
Control group (combined)			
Pre			
x̄	85.2	80.9	79.4
SD	3.4	10.4	20.9
Post			
x̄	84.3	77.7	81.9
SD	7.4	12.3	12.3
Change			
x̄	−0.9	−3.2	2.5
SD	6.2	15.5	14.4

EXP group, n = 15; CONTROL group, n = 16.

TABLE 4. *Means, standard deviations, and change score means for timing focus index (%)*

	Gastrocnemius	Tibialis anterior	Vastus lateralis
Experimental group			
Pre			
x̄	107.0	82.6	77.3
SD	23.5	13.1	9.7
Post			
x̄	100.1	88.5	86.4
SD	24.7	14.5	14.2
Change			
x̄	−6.9	6.0	9.1
SD	14.0	12.3	13.9
Control groups (combined)			
Pre			
x̄	97.1	92.4	97.0
SD	23.7	14.8	20.9
Post			
x̄	101.5	89.7	90.5
SD	30.8	20.5	18.1
Change			
x̄	4.3	−2.7	−6.4
SD	17.8	12.8	14.2

EXP group, n = 15; CONTROL group, n = 16.

the GA and TA muscle increased by 13%, the latter approaching a statistically significant difference in comparison with control group changes ($F = 4.32$; $p = 0.0565$). No other significant changes were seen in the symmetry analysis.

A timing index was computed to reflect the temporal focus of EMG activity without identifying onset and termination times. A large value would indicate EMG activity focused in a small percentage of the gait cycle, whereas a small value would reflect EMG activity spread throughout the gait cycle. The data for each muscle group are listed in Table 4. Statistically significant differences between EX and control group, indicating a more focused activation period, were seen only in the VL muscle ($F = 6.43$; $p = 0.0220$).

Pretest to posttest changes in the activation periods of the GA, TA, and VL muscles in the EX group were further investigated by determination of onset and termination times of EMG. Significant changes were observed in the durations of the VL ($t = 2.59$; $p = 0.0303$) and TA muscle ($t = 2.30$; $p = 0.0471$; Table 5). The duration of both muscle groups was shortened almost entirely by a quicker termination, ~4% GC for VL and 6% for TA. The onsets were as expected, during the preswing phase, at ~88% GC for VL and 58% for TA. The termination times were later than normal during pretest (17). TA termination is normally before the end of the loading phase, at 9% GC. Pretest TA termination was at 14% GC with posttest termination at 8% GC. The VL termination is normally in early mid-stance ~30% GC. Pretest VL termination was at 38.5%, and posttest termination, at 34.5% GC.

TABLE 5. *Means, standard deviations, and change score means for EMG onset and termination (% GC)*

Experimental group	Onset	Termination	Duration
Tibialis anterior			
Pre			
x̄	58.7	14.4	
SD	3.4	10.5	
Post			
x̄	57.6	8.0	
SD	3.2	7.9	
Change			
x̄	−1.1	−6.5	−5.4
SD	2.2	6.9	7.4
Vastus lateralis			
Pre			
x̄	87.7	38.5	
SD	4.0	6.3	
Post			
x̄	87.7	34.5	
SD	3.5	7.7	
Change			
x̄	0.1	−4.0	−4.0
SD	2.5	5.9	5.0

EXP group, n = 15; CONTROL group; n = 16.

DISCUSSION

Stride Parameters

The major goal of RAS training was to increase the normal gait velocity of PD patients for flat surface and incline walking. Subjects who participated in the RAS training improved on average 25% in both of these tasks. The improvement was facilitated by nearly equal percentage increases in cadence and stride length. The control group that participated in self-paced training also improved their gait velocity, but by less than one third of the improvement seen in the RAS group. This control group increased velocity solely with an increase in stride length but no apparent change in cadence. Subjects in the no-training control group showed no improvement, as would be expected from such a control condition. The difference in increase in velocity in the EX group over the SPT group appeared to be facilitated by the addition of RAS to the exercise program. In only 3 weeks, the RAS program, which was based on increasing walking tempo gradually through rhythmic auditory cuing, was effective in shifting the subjects' intrinsic gait tempo. Part of the mechanism involved with this tempo shift may have been a rhythmic entrainment effect evidenced by the subject's ability to reproduce the tempo of the musical rhythm without cuing. Although motivational factors through the music cannot be excluded as a reason for enhanced gait performance, their effects were minimized by the fact that each subject had to train with the same musical selection for 3 weeks. Repeated use of music of relatively low complexity is assumed to induce a great amount of redundancy into the perceptual process and thus strongly reduce affective arousal effects related to motivation (18). This assumption was supported by comments from the study subjects that the music seemed to get repetitive over time.

EMG Patterns

Some differences in variability, symmetry, and timing of EMG measures were observed between EX and control group data. However, the observed changes were fairly small and not consistent across muscles. Data pooling of the two control groups may have also masked EMG changes between NT and SPT subjects. However, similar trends in timing and variability of EMG patterns under the influence of auditory rhythm have been reported in previous studies (7,8,9,19,20).

The slight improvements in EMG symmetry tended to be actually more dramatic in subjects who had poor pretest symmetry, whereas subjects with good pretest symmetry generally improved little. This was illustrated by the decrease in standard deviation across the subject pool of the pretest and posttest symmetries. Pretest standard deviation values were 24.3% for GA and 16.5% for TA; posttest values were 8.2% and 7.7%, respectively.

Significant changes were also seen in the VL muscle activation period, with activity focused in a smaller time percentage of the gait cycle. A further analysis of EMG onset and termination times revealed that the VL and TA muscles had EMG onset times that corresponded to the proper phase of the gait cycle in both pretest and posttest. However, termination time showed a slight delay in the pretest trials. Biomechanical analysis would be needed to show if the mid and terminal stance periods had changed, or if the late EMG activity represented nonfunctional rigidity. Either way, at posttest, both termination times had shifted to a more normal profile.

Mechanisms

The subjects' ability to reproduce without cuing the fastest training cadence 24 h after the last training session indicates the possible effect of rhythmic entrainment mechanisms. The small error in reproduction was especially remarkable in light of previous findings that show that time estimation, recall, and reproduction in PD patients are impaired (11). Auditory rhythm may have acted as an external timekeeper clock to which the step cadence became synchronized during the training phase, thus helping to stabilize destabilized internal time keeping and rhythm formation processes in PD patients (10,11). Several patients reported pacing themselves by singing the music silently.

In summary, RAS training improved gait velocity, cadence, and stride length significantly after only 3 weeks. In addition, some features of EMG gait-cycle profiles changed toward more normal muscle-activation patterns. The data suggest a viable role for RAS as a sensorimotor-based technique for gait facilitation in PD patients.

Acknowledgments: This research was sponsored by a grant from the U.S. Department of Health and Human Services, Administration on Aging, 90AM1684-01.

REFERENCES

1. Knutsson E. An analysis of parkinsonian gait. *Brain* 1972; 95:475–486.

2. Giladi N, McMahon D, Przedborski S, et al. Motor blocks in Parkinson's disease. *Neurology* 1992;42:333–339.
3. Weiner WJ, Singer C. Parkinson's disease and nonpharmacologic treatment programs. *J Am Geriatr Soc* 1989;37:359–363.
4. Caird FI. *Rehabilitation in Parkinson's disease*. New York: Chapman & Hall, 1991.
5. Martin JP. *The basal ganglia*. Toronto: JB Lippincott, 1967.
6. McIntosh GC, Thaut MH, Rice RR, Miller RA. Stride frequency modulation in parkinsonian gait using rhythmic auditory stimulation. *Ann Neurol* 1994;36:316.
7. Richards CL, Malouin F, Bedard PJ, Cioni M. Changes induced by L-dopa and sensory cues on the gait of parkinsonian patients. In: Woollacott M, Horak F, eds. *Posture and gait: control mechanisms*. Eugene, OR: University of Oregon Books, 1992:126–129.
8. Thaut MH, McIntosh GC, Prassas SG, Rice RR. Effect of rhythmic auditory cuing on temporal stride parameters and EMG patterns in normal gait. *J Neurol Rehabil* 1992;6:185–190.
9. Thaut MH, McIntosh GC, Prassas SG, Rice RR. Effect of rhythmic auditory cuing on temporal stride parameters and EMG patterns in hemiplegic gait of stroke patients. *J Neurol Rehabil* 1993;7:9–16.
10. Freeman JS, Cody FWJ, Schady W. The influence of external timing cues upon the rhythm of voluntary movements in Parkinson's disease. *J Neurol Neurosurg Psychiatry* 1993;56:1078–1084.
11. Pastor MA, Artieda M, Jahanshahi M, Obeso JA. Time estimation and reproduction is abnormal in Parkinson's disease. *Brain* 1992;115:211–225.
12. Thaut MH, Rathbun J, Miller RA. Music vs metronome timekeeper in a rhythmic motor task. (submitted for publication.)
13. Winter DA, Yack HJ. EMG profiles during normal human walking: stride-to-stride and inter-subject variability. *Electroencephalogr Clin Neurophysiol* 1987;67:402–411.
14. Arsenault AB, Winter DA, Marteniuk RG. Bilateralism of EMG profiles in human locomotion. *Am J Phys Med* 1986;65(1):1–15.
15. Oeberg T, Karsznia A, Oeberg K. Basic gait parameters: reference data for normal subjects, 10–79 years of age. *J Rehabil Res Dev* 1993;30:210–223.
16. Miller RA, Thaut MH, McIntosh GC, Rice RR. Components of EMG symmetry and variability in parkinsonian and healthy elderly gait. *Electroencephalogr Clin Neurophysiol* (in press).
17. Perry J. *Gait analysis: normal and pathological function*. Thorofare, NJ: Slack, 1992.
18. Berlyne D. *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts, 1971.
19. Rossignol S, Melvill Jones G. Audio-spinal influences in man studied by the H-reflex and its possible role on rhythmic movements synchronized to sound. *Electroencephalogr Clin Neurophysiol* 1976;41:83–92.
20. Safranek M, Koshland G, Raymond G. Effect of auditory rhythm on muscle activity. *Phys Ther* 1982;62:161–168.

ELSEVIER

Journal of Neurological Sciences 151 (1997) 207–212

JOURNAL OF THE NEUROLOGICAL SCIENCES

Rhythmic facilitation of gait training in hemiparetic stroke rehabilitation

M.H. Thaut[a,*], G.C. McIntosh[a,b], R.R. Rice[b]

[a]Center for Research in Neurorehabilitation, Department of Electrical Engineering and Department of Music, Colorado State University, Fort Collins, CO 80 523, USA
[b]Center for Research in Neurorehabilitation, Poudre Valley Hospital, Fort Collins, CO 80 524, USA

Received 29 August 1996; revised 31 March 1997; accepted 19 April 1997

Abstract

Experimental and control groups of 10 hemiparetic stroke patients each underwent a 6 week, twice daily gait training program. The control group participated in a conventional physical therapy gait program. The experimental group trained in the same basic program with the addition of rhythmic auditory stimulation (RAS). Patients entered the study as soon as they could complete 5 strides with hand-held assistance. The training program had to be completed within 3 months of the patients' stroke. In the experimental group RAS was used as a timekeeper to synchronize step patterns and gradually entrain higher stride frequencies. Study groups were equated by gender, lesion site, and age. Motor function was assessed at pretest using Barthel, Fugl-Meyer, and Berg Scales. Walking patterns were assessed during pre- and post-test without RAS present. Pre- vs post-test measures revealed a statistically significant ($P<0.05$) increase in velocity (164% vs 107%), stride length (88% vs 34%), and reduction in EMG amplitude variability of the gastrocnemius muscle (69% vs 33%) for the RAS-training group compared to the control group. The difference in stride symmetry improvement (32% in the RAS-group vs 16% in the control group) was statistically not significant. The data offer evidence that RAS is an efficient tool to enhance efforts in gait rehabilitation with acute stroke patients. © 1997 Elsevier Science B.V.

Keywords: Stroke; Gait; Rehabilitation; Auditory rhythm

1. Introduction

Motor dysfunction is one of the most frequently encountered and therapeutically persistent problems after stroke. Therefore, recovery of motor function is a major emphasis in almost all rehabilitation efforts for stroke patients. Motor deficit characterized by hemiparesis is a common manifestation of cerebral hemispheric stroke in the middle cerebral artery vascular distribution. One of the most desired outcomes of rehabilitation is the improvement of ambulatory function since it determines to a large degree the status of the patient in respect to activities of daily living and associated quality of life (Richards et al., 1993). Current programs in stroke rehabilitation have met with varying success. A recent assessment of the efficacy of stroke rehabilitation shows mixed results (Jeffery and Good, 1995). For example, a study by Hesse et al. (1994) with mildly affected stroke patients who were mostly past the acute recovery stage of 3 months showed significant improvement in gait velocity and aspects of stride symmetry, yet endurance, symmetry of ground-reaction forces and functional performance did not improve after daily training for 4 weeks. Therefore, the further refinement of efficient rehabilitation techniques remains an important challenge.

Gait of hemispheric stroke patients is characterized by several abnormal features. Among those features are varied degrees of asymmetry in stride times and stride length, slowed velocity, poor joint and posture control, muscle weakness, abnormal muscle tone, and abnormal muscle activation patterns, mostly affecting the paretic side. What is important to note is that the resulting deficits in gait

*Corresponding author. Tel: +1 970 4917384; fax: +1 970 4917541; e-mail: mthaut@lamar.colostate.edu

0022-510X/97/$17.00
PII S0022-510X(97)00146-9

performance are not only due to muscle weakness but to complex abnormalities in motor control (Good, 1994). Therefore, rehabilitation strategies for gait recovery need to address the facilitation of appropriate motor control strategies.

The recent emergence of new data regarding the physiologic mechanisms underlying recovery, specifically in respect to the facilitation of cortical reorganization and the application of learning and training paradigms, may provide new avenues to develop strategies to enhance motor recovery. For instance, several authors have noted evidence for the possible stimulation or 'unmasking' of intact alternate motor control centers (Fries et al., 1993; Chollet et al., 1991; Brion et al., 1989; Freund and Hummelsheim, 1985), and the modulation of cortical motor output through motor training (Pascual-Leone et al., 1993; Aizawa et al., 1991; Bach-y-Rita, 1992). Research data are yet equivocal as to what extent the utilization of sensorimotor facilitation can strengthen training paradigms and/or shift motor control strategies in motor rehabilitation (Good, 1994). However, recent data suggest rehabilitative procedures that involve highly repetitive, rhythmically patterned movement training to be particularly effective (Buetefisch et al., 1995), possibly by facilitating long-term potentiation in the sensorimotor cortex as a mechanism for motor learning (Asanuma and Keller, 1991a,b).

Therefore, to further clarify the role of sensory stimuli in motor recovery we sought to determine the usefulness of auditory rhythm as an external timekeeper to enhance efforts in gait rehabilitation with stroke patients within 3 months post cardiovascular accident (CVA). Our study was based on previous work in which we had used auditory rhythm in an entrainment design to study the immediate effect on gait patterns in stroke patients without training effect (Thaut et al., 1993). Our data showed a significant improvement in motor unit recruitment patterns, weight bearing stance time on the paretic leg, and stride symmetry in 10 hemispheric stroke patients, ranging from 4 weeks to 2 years post stroke. These results were repeated in 3 consecutive trials, each 2 weeks apart from each other. In the current study we investigated the rhythmic entrainment effect as a therapeutic technique in gait training of acute stroke patients within 3 weeks post CVA.

2. Methodology

2.1. Subjects

Twenty subjects, 10 male and 10 female, were randomly assigned to either an experimental group, using rhythmic auditory stimulation (RAS) with conventional physical therapy (PT), or a control group, using only conventional PT for gait training. Conventional PT was based on the Neurodevelopmental Treatment (NDT) approach. Each group was matched by gender (5 male and 5 female patients), and lesion site (5 right- and 5 left hemispheric strokes; localized by MRI scan). The mean age was 73±7 for the RAS-group, and 72±8 for the control group. The 2 groups were further assessed at the outset of the experiment by a physical therapist blind to the experiment on the Barthel Index (Mahoney and Barthel, 1954), the Fugl-Meyer Scale (Fugl-Meyer et al., 1975) and the Berg Scale (Berg et al., 1989). The Barthel Index was also readministered at the post-test. The Barthel Index for the RAS-group was 53 at pre-test and 86 at post-test, and 50 and 82 respectively for the control group. The Fugl-Meyer score was 9 for balance and 25 for lower extremity function in both groups. The Berg score was 45 for the RAS-group and 50 for the control group.

In the RAS-group, 5 subjects had right-hemispheric middle cerebral artery (MCA) strokes, 3 had left internal capsule (IC) and 2 had left MCA strokes. Three subjects had suffered a second strokes. Six strokes were ischemic and 4 hemorrhagic.

In the control group, 4 subjects had right MCA strokes, 1 had a right IC stroke, 2 had left MCA and 3 left IC strokes. Two subjects had suffered a second stroke. Five strokes were hemorrhagic and 5 ischemic.

Mild to moderate distal sensory dysfunction was manifested in all MCA distribution strokes and not displayed in the patients with IC strokes. Both groups had lower limb spasticity, mostly seen in knee flexors/extensors, plantar flexor, and hip flexion/extension patterns, as typical for stage 4 of the hemiplegia recovery scale described by Brunnstrom (1970).

2.2. Training

Patients entered the study as soon as they could complete 5 strides with hand-held assistance by the therapist, i.e. supporting the forearm, wrist and elbow at approximately 90 degree elbow flexion on the nonparetic side, within 3 weeks post CVA. Hand-held assistance was also given to the patient throughout experimental trials and gait training time when needed. The average entry post CVA for the study was 16.1±4 days for the RAS-group and 15.7±4 days for the control group. The entire training duration for the study was 6 weeks. In both groups patients trained twice daily, 30 min each in the morning and the afternoon 5 days a week. A pool of 4 physical therapists, 2 for each group, was specifically trained to handle patients in both groups in order to warrant a maximum degree of consistency in training style. Patients who were in separate treatment groups were also assigned separate rooms in the hospital for the duration of the study. In both groups the total walking time and distance was tracked to ensure that both groups exercised approximately the same amount. Pre-gait exercises were not included in the 60 min training time in this study, and were carried out in similar fashion in both groups if therapeutically indicated.

In the RAS-group, patients trained their walking by using a metronome or specifically prepared music tapes. After an initial cadence assessment during a 1 to 2 min warm-up walk, the rhythm frequency was matched to the gait cadence for the first quarter of the session.

During the second and third quarter the rhythm frequency was incrementally increased by 5 to 10%, depending on the patient's ability. The last quarter was spent with RAS intermittently faded to train for independent carry-over of improved gait patterns. The control group trained the same amount of time and distance with equivalent instructions regarding speed improvement, however, without RAS facilitation.

2.3. Testing

All patients were pre- and post-tested the day before commencing and the day after concluding the training. Pre- and post-tests were carried out without RAS. For testing patients walked along an 10 meter flat walkway with data recorded only along the middle 6 meters. Stride timing was recorded at a sampling rate of 500 per s with a computerized foot sensor system consisting of 4 foot contact sensors (heel, 1st and 5th metatarsal, big toe) embedded into shoe inserts, a portable microprocessor to record data, and computer interface and data analysis hard and software. Electromyographic activity (EMG) of the medial gastrocnemius (MG) was recorded with surface electrodes; bipolar silver/silver chloride, 8 mm diameter with 2 cm spacing embedded in a plastic enclosure. On-site preamplification of 35 (v/v) was followed by an amplification of 10,000 to 50,000 with highpass filtering at 20 Hz before recording at 500 samples per s with a 12 bit analog to digital converter. Electrode placement, longitudinally along the major belly portion of the muscle, was performed by a physical therapist experienced in electromyography.

2.4. Data analysis

Stride parameters of 5 stride cycles were used to assess improvement in gait ability with regard to velocity, stride length, and swing symmetry. Symmetry was calculated as the time ratio between the swing times of 2 successive steps using the longer step as the denominator. Percentage change scores for all stride data were computed for each subject and averaged across groups for statistical analysis. Percentage change scores were chosen to offset individual performance differences at the outset of the study. Mann-Whitney rank-order tests were used for statistical analysis of differences between groups. Nonparametric statistics were selected to offset possible violations of normalcy in percentage score distributions in relatively small samples.

Variability of EMG shape patterns was computed on the full-wave RMS-rectified signal of the paretic leg by calculating integrated amplitude values (in μV) and amplitude ratios (in μV/ms) and their respective standard deviations. Five strides were used to compute an ensemble average of the shape of the EMG curve (12).

2.5. RAS

The rhythmic stimulus in the training sessions consisted of music tapes played over headsets that were prerecorded on a synthesizer/sequencer module. The module was used to record the same music digitally at various frequencies suitable for the patients' gait cadence. The sequencer was used as a variable frequency driver for the music. Instrumental music in 4 different styles was prepared (classic, folk, country, jazz). The music was recorded in 2/4 meter to match the rhythm of the step patterns in gait. A metronome beat was overlaid on the strong beat of the music to enhance the rhythmic perception for the patient. Rhythmic and melodic patterns in-between the metronome beats subdivided the basic meter in ratios of 1:2 and 1:4.

3. Results

3.1. Stride parameters

During pre-test both groups showed highly abnormal stride data compared to normal age-matched data (Oeberg et al., 1993) which are reported in the literature as 73 m/min for velocity, 1.27 m for stride length, and 113 steps/min for cadence. The mean velocity for the RAS-group was 19.7±11 m/min and 17.3±7 m/min for the control group. Stride length was shortened to 0.64±0.31 m for the RAS-group and 0.55±0.11 m for the control group. Strong lower-limb hemiparesis was evidenced by a mean swing symmetry ratio of 0.64±0.16 for the RAS-group and 0.61±0.25 for the control group. Mean cadence values were 63±10 steps/min for the RAS group, and 62±20 steps/min for the control group.

Post-test data show that both groups improved their stride parameters over the 6-week therapy period. However, there were significant differences in the recovery rate between experimental and control conditions. The mean velocity had increased in the RAS-group to 48±18 m/min, and to 32±10 m/min in the control group. RAS-trained subjects lengthened their stride on the average to 1.00±0.30 m, and control subjects had increased their stride length to 0.69±0.19 m. Mean symmetry ratios improved considerably for the RAS-group (0.82), and to a

Table 1
Results

	Velocity (m/m)	Stride length (m)	Symmetry	Cadence
RAS	19.7±11	0.64±0.31	0.64±0.16	63±10
group	48.0±18	1.00±0.30	0.82±0.14	98±17
Control	17.3±7	0.55±0.11	0.61±0.25	62±20
group	32.0±10	0.69±0.19	0.68±0.23	90±16

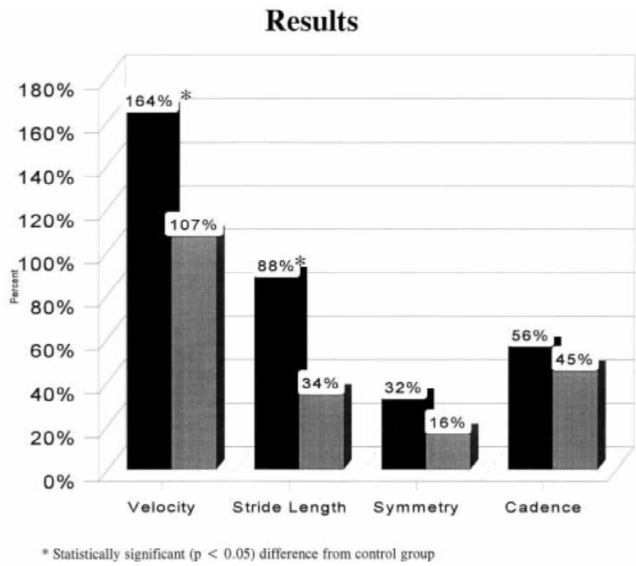


Fig. 1. Percentage change scores from pre- and post-test between RAS and control group for velocity, stride length, symmetry, and step cadence.

lesser degree for the control group (0.68). Mean cadence improved to 98±17 steps/min in the RAS-group, and to 90±16 steps/min in the control group (Table 1).

Percentage change scores for all stride data are summarized in Fig. 1. The mean percentage increase in velocity was 164% for the RAS-group, and 107% for the control group. The difference between groups was statistically significant (*MW* [*MW*=Mann Whitney]$_{calc}$=132, $P<0.05$). Stride length had improved by 88% in the RAS-group compared to 34% for the controls. Group differences for stride length were also statistically significant (MW_{calc}=136, $P<0.02$). Although RAS-training had improved symmetry by 32% compared to 16% for conventional PT, the differences in improvement rate between the 2 groups were statistically not significant (MW_{calc}=129, $P=0.09$). Likewise, differences in step cadence improvement between the 2 groups (56% for RAS, 45% for PT) were nonsignificant, indicating that the velocity improvement in the RAS-group was mostly caused by increased stride length.

3.2. EMG analysis

A variability analysis was computed on the ensemble average of the shape of the EMG-curve of the gastrocnemius muscle on the paretic limb, using the standard deviations of the integrated amplitude values. Coefficients of variation (Standard Deviation/Mean x 100) were computed to normalize variability ratios. In the RAS-group, coefficients of variations were 69±11% lower at post-test than at pretest. In the control group coefficients of variations had decreased from pre- to post-test by 33±31%. Group differences for EMG variability were statistically significant (MW_{calc}=138; $P<0.02$). Sample traces from pre- and post-test recordings from a RAS-group subject are given in Fig. 2.

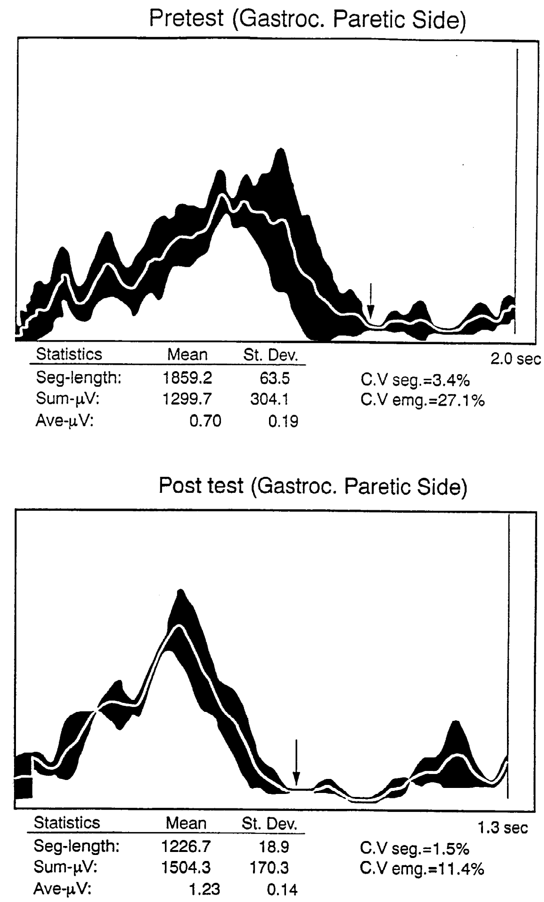


Fig. 2. Averaged RMS full-wave rectified EMG traces (5 step cycles) from a RAS subject. Muscle shown is the gastrocnemius on the paretic side (right). Shaded area around the EMG curve is 1 standard deviation band around the mean. Arrow denotes toe-off. Graph begins and ends on heelstrike. Statistics under the graph show segment length (in ms), sum and averaged amplitude (in μV), and the coefficients of variation of averaged segment length and EMG amplitude. Clearly visible is the reduction in amplitude variability from pre- to post-test.

4. Discussion

Pre- to post-test comparisons between 2 closely matched groups of stroke patients showed that rhythmic facilitation of gait training significantly improved gait velocity and stride length relative to gait training without rhythmic facilitation. Rhythmic facilitation also produced a noticeable improvement in stride symmetry compared to the control group. However, the difference between the 2 groups was not significantly different. Whereas velocity increases in the RAS-group were mainly driven by lengthening of strides (88%) and to a smaller degree by faster step rates (56%), those contributions were reversed in the control group, where increases in step frequency were higher (45%) than changes in stride length (34%).

Obviously, velocity increases which are proportionally driven by more step lengthening than higher step frequencies result in a more efficient gait pattern. The reason that the velocity increase of the control patients relied more on higher step rates may have been due to the fact that improvements in stride length were compromised in the control group by the persistence of high asymmetry in the step patterns.

An important finding for gait recovery is the large degree of restoration of swing symmetry after the RAS-training. Asymmetry is a persistent feature in gait patterns of stroke patients, and is very resistant to rehabilitation efforts. Improved symmetry allows for more normal gait patterns, higher velocity, and more evenly distributed exercise of both lower limbs. The timing symmetry inherent in the rhythmic signal may have served as an efficient cue for the patient to achieve a higher degree of temporal stride symmetry accompanied by lengthened stride. Similar to Hesse et al. (1994), NDT provided a more limited restoration of gait symmetry in our study. Gait facilitation through other sensorimotor systems using, e.g. visual cuing of stride length, to improve gait symmetry has been successfully demonstrated by Montoya et al. (1994). However, whereas visual cuing may preferentially access spatial control parameters of movement, Richards et al. (1992) have proposed that RAS may act on more central facilitation mechanisms since the symmetry of stride times as well as stride length have been shown to improve with RAS (Prassas et al., 1997). However, the lack of statistical significance of a relatively large group difference in our study sample due to inconsistency in improvement patterns across patients underlines the problems in finding effective methods to restore gait symmetry in hemiparetic stroke patients.

An interesting finding which has been noticed in previous research (Thaut et al., 1993; Miller et al., 1996; Rossignol and Melvill Jones, 1976; Safranek et al., 1982) is the effect of auditory rhythm on EMG activity, especially the reduction of amplitude variability as a result of rhythmic training. Our data suggest that auditory rhythmic timekeepers may enhance more regular motor unit recruitment patterns. The functional significance of this effect for motor control is not entirely clear. However, it underscores the existence of physiological mechanisms between the auditory and the motor systems. The ability of auditory rhythm to effectively entrain motor patterns and also influence nontemporal parameters such as stride length, may help to assign rhythmic auditory stimuli a larger role in motor control than previously assumed. Considering (a) the particular effectiveness of the auditory system to process timing information with a high degree of speed and accuracy, and (b) the fundamental importance of timing for all parameters of complex movement, a model of rhythmic auditory-motor entrainment provides an intriguing context for further study. Hemispheric stroke patients may particularly benefit from RAS since auditory rhythm is processed bilaterally and no difference in performance was observed in this study as in our previous work (Thaut et al., 1993) between left- and right-hemispheric patients.

To control for motivational–emotional factors in the music to enhance gait performance, the same music was used for a patient's training period. Repetitive use of music of relatively low complexity has been shown to provide redundancy in music perception which strongly reduces affective arousal related to motivational states (Berlyne, 1971).

Substantial increases in walking speed are an important functional goal in gait rehabilitation if safe gait patterns, e.g., avoiding an increased risk of falling, are maintained. In our study sample, increased velocity was accompanied by substantial increases in stride symmetry which may be considered one indicator that postural stability was not compromised. Functional benefits of higher walking speed include reduced travel time which may help to increase the ambulatory range for the patient as well as reduce physical fatigue. In the current state of gait rehabilitation, techniques that complement and enhance therapeutic efforts are greatly needed. This study provided the first evidence that rhythmic entrainment mechanisms, utilized as a training device for stroke patients, can improve important gait parameters within a rehabilitative context.

Acknowledgments

This research was funded in part by a grant from the Poudre Valley Hospital Foundation and grants RR 07127-20 and RR 07127-23 from the National Institutes of Health (NIH).

References

Aizawa, H., Inase, M., Mushiake, H., Shima, K., Tanji, J., 1991. Reorganization of activity in the supplementary motor area associated with motor learning and functional recovery. Exp. Brain Res. 84, 668–671.

Asanuma, H., Keller, A., 1991. Neuronal mechanisms of motor learning in mammals. Neuroreport 2, 217–224.

Asanuma, H., Keller, A., 1991. Neurobiological basis of motor learning and memory. Conc. Neurosci. 2, 1–30.

Bach-y-Rita, P., 1992. Recovery from brain damage. J. Neurol. Rehabil. 6, 101–200.

Berg, K., Wood-Dauphine, S., William, J.I., 1989. Measuring balance in the elderly: preliminary development of an instrument. Physiother. Can. 41, 304–311.

Berlyne, D., 1971. Aesthetics and Psychobiology. Appleton-Century-Crofts, New York.

Brion, J., Demeurisse, G., Capon, A., 1989. Evidence of cortical reorganization in hemiparetic patients. Stroke 20, 1079–1084.

Brunnstrom, S., 1970. Movement therapy in hemiplegia. Harper & Row, New York.

Buetefisch, C., Hummelsheim, H., Denzler, P., Mauritz, K.H., 1995. Repetitive training of isolated movements improves the outcome of motor rehabilitation of the centrally paretic hand. J. Neurol. Sci. 130, 59–68.

Chollet, F., DiPiero, V., Wise, R.J.S., Brooks, D.J., Dolan, R.J., Frackowiak, R.S.J., 1991. The functional anatomy of motor recovery after stroke in humans: a study of positron emission tomography. Ann. Neurol. 29, 63–71.

Freund, H.J., Hummelsheim, H., 1985. Lesions of premotor cortex in man. Brain 108, 697–733.

Fries, W., Danek, K., Scheidtmann, K., Hamburger, C., 1993. Motor recovery following stroke, Role of descending pathways from multiple motor areas, Brain 116, pp. 369–382.

Fugl-Meyer, A.R., Jaasko, L., Leyman, L., Olsson, S., Steglind, S., 1975. The post-stroke hemiplegic patient. I. A method for evaluation of physical performance. Scand. J. Rehab. Med. 7, 13–31.

Good, D.C., 1994. Treatment strategies for enhancing motor recovery in stroke rehabilitation. J. Neurol. Rehabil. 8, 177–186.

Hesse, S.A., Jahnke, M.T., Bertelt, C.M., Schreiner, C., Luecke, D., Mauritz, K.H., 1994. Gait outcome in ambulatory hemiparetic patients after a 4 week comprehensive rehabili tation program and prognostic factors. Stroke 25, 1999–2004.

Jeffery, D.R., Good, D.C., 1995. Rehabilitation of the stroke patient. Curr. Opin. Neurol. 8, 62–68.

Mahoney, F.D., Barthel, B.W., 1954. Rehabilitation of the hemiplegic patient: a clinical evaluation. Arch. Phys. Med. Rehabil. 35, 359–362.

Miller, R.A., Thaut, M.H., McIntosh, G.C., Rice, R.R., 1996. Components of EMG symmetry and variability in Parkinsonian and healthy elderly gait. Electrencephal. Clin. Neurophysiol. 101, 1–7.

Montoya, R., Dupui, P., Pages, B., Bessou, P., 1994. Step-length biofeedback device for walk rehabilitation. Med. Biol. Eng. Comput. 32, 416–420.

Oeberg, T., Karsznia, A., Oeberg, K., 1993. Basic gait parameters: Reference data for normal subjects, 10–79 years of age. J. Rehabil. Res. Develop. 30, 210–223.

Pascual-Leone, A., Cohen, L.G., Dang, N., Brasil-Neto, J.P., Cammarota, A., Hallett, M., 1993. Acquisition of fine motor skills in humans is associated with the modulation of cortical motor outputs. Neurology 43,(suppl 2) A157.

Prassas S.G., Thaut, M.H., McIntosh, G.C., Rice, R.R., 1997. Effect of auditory rhythmic cuing on gait kinematic parameters in hemiparetic stroke patients. Gait Posture, in press.

Richards, C.L., Malouin, F., Berard, P.J., Cioni, M., 1992. Changes induced by L-dopa and sensory cues on the gait of parkinsonian patients. In: Woollacott, M., Horak, F. (Eds.), Posture and Gait: Control Mechanisms. University of Oregon Books, Eugene OR, pp. 126–129.

Richards, C.L., Malouin, F., Wood-Dauphine, S., Williams, J.I., Bouchard, J.-P., Brunet, D., 1993. Task-specific physical therapy for optimization of gait recovery in acute stroke patients. Arch. Phys. Med. Rehabil. 74, 612–620.

Rossignol, S., Melvill Jones, G., 1976. Audio-spinal influences in man studied by the H-reflex and its possible role in rhythmic movement synchronized to sound. Electroencephal. Clin. Neurophysiol. 41, 83–92.

Safranek, J.G., Koshland, G., Raymond, G., 1982. Effect of auditory rhythm on muscle activity. Phys. Ther. 62, 161–168.

Thaut, M.H., McIntosh, G.C., Rice, R.R., Prassas, S.G., 1993. Effect of rhythmic auditory cuing on EMG and temporal stride parameters and EMG patterns in hemiparetic gait of stroke patients. J. Neurol. Rehabil. 7, 9–16.

PERGAMON

Neuropsychologia 40 (2002) 1073–1081

NEUROPSYCHOLOGIA

www.elsevier.com/locate/neuropsychologia

Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients

M.H. Thaut [a,*], G.P. Kenyon [a], C.P. Hurt [a], G.C. McIntosh [a], V. Hoemberg [b]

[a] *Center for NeuroRehabilitation Research, Colorado State University, Fort Collins, CO 80523, USA*

[b] *Neurological Therapy Center, Heinrich-Heine University Duesseldorf, 40591 Duesseldorf, Germany*

Received 9 November 1999; received in revised form 9 April 2001; accepted 8 June 2001

Abstract

The effect of rhythmic cueing on spatiotemporal control of sequential reaching movements of the paretic arm was studied in 21 hemispheric stroke patients. Reaching movements were studied with and without rhythmic metronome cuing in a counterbalanced design. Metronome frequencies were entrained to the naturally selected frequency of the patient. Results indicate statistically significant ($P < 0.05$) improvements of spatiotemporal arm control during rhythmic entrainment. Variability of timing and reaching trajectories were reduced significantly. Time series analysis of sequential movement repetitions showed an immediate reduction in variability of arm kinematics during rhythmic entrainment within the first two to three repetitions of each trial. Rhythm also produced significant increases in angle ranges of elbow motion ($P < 0.05$). Analysis of acceleration and velocity profiles of the wrist joint showed significant kinematic smoothing during rhythmic cuing. The link between rhythmic sensory timing and spatiotemporal motor control was investigated using a mathematical optimization model with minimization of peak acceleration as criterion. Rhythmically cued acceleration profiles fit the predicted model data significantly closer ($P < 0.01$) than the self-paced profiles. Since velocity and acceleration are mathematical derivatives of position–time trajectories, the model data suggest that enhanced timing precision via temporal phase and period coupling of the motor pattern to the rhythmic time timekeeper enhances the brain's computational ability to optimally scale movement parameters across time.

Keywords: Stroke; Arm kinematics; Rehabilitation; Motor learning; Mathematical modeling

1. Introduction

Considering the high incidence of motor impairments associated with stroke [4,36], recovery of motor function is a central part of neurologic rehabilitation efforts for patients with stroke. Approximately 70–88% of persons with ischemic hemispheric strokes have some degree of motor impairment which is frequently manifested in hemiplegia of the lower face, arm, and leg opposite to the site of the infarction [13]. Analysis of affected motor functions has shown that impairment of arm function is more common than leg impairment, and is also more resistant to rehabilitation efforts [23,36]. Several conventional therapy program for arm therapy are available in neurologic rehabilitation following the concepts of, e.g. Brunnstrom, the Bobath or neurodevelopmental treatment approach, or proprioceptive neuromuscular facilitation (see for a review [31]). However, outcome research has shown equivocal results and scientific rationales for these approaches have come increasingly under criticism [6,13,10]. Thus, the continuing need persists for research regarding the neurological mechanisms of motor control and therapeutic strategies to restore motor function [3].

New approaches which utilize sensorimotor systems in learning and training designs to restore motor function are beginning to show promising results in rehabilitation research [19]. One such approach, a rhythmic model of rehabilitative motor training, has shown significant improvements in gait function of stroke patients [24,39,41]. In this model, rhythm functions as a sensory cue to induce temporal stability and enhance the temporal organization of motor control in the nervous system by translating the temporal structure of movement patterns into temporally isomorphic auditory rhythmic patterns to entrain the movement in question. Similar models have been successfully used in high-performance motor skill learning in sports and music. In gait training, based on models of limit cycle entrainment of coupled harmonic oscillators [24,41], frequencies of auditory rhythmic stimuli (using symmetric pulse patterns in metronome or embedded in musical patterns) were matched

* Corresponding author. Tel.: +1-970-491-7384; fax: +1-970-491-7541.
E-mail address: mthaut@lamar.colostate.edu (M.H. Thaut).

0028-3932/02/$ – see front matter
PII: S0028-3932(01)00141-5

to the patients' baseline frequency to initially entrain the current 'natural frequency' and stabilize gait parameters. Subsequent step-wise entrainment to higher step frequencies resulted in comprehensive improvements in the stability and variability of gait kinematic parameters as well as amplitude and time variability of muscle activation patterns. These gains were significantly higher in controlled clinical studies in comparison to other rehabilitation approaches [39].

Based on the previously noted effect of rhythmic facilitation on hemiparetic gait, we sought to investigate the effect of rhythm on the control of paretic arm movements in stroke patients. Unlike gait patterns, which are rhythmic in nature and thought to be controlled by physiological pattern generators [14], most functional arm movements are discrete, biologically nonrhythmic, and volitional. However, skill learning of arm and hand functions in high performance environments, such as sports or music employs very successfully rhythmical stimuli to pattern and cue the programming and execution of motor skills. Also, recent approaches in neurologic motor rehabilitation, emphasizing repetitive training paradigms [6,15,18], have shown beneficial results and are putatively supported by neurophysiological research regarding long-term potentiation, structural transformation of synapses, and modulation of cortical motor output [1,20,22,28,34].

The purpose of this study, therefore, was two-fold. In order to assess the effect of rhythm on paretic arm function, we employed an isochronous metronome stimulus to cue reaching movements in arm training of hemiparetic stroke patients, and compared the data to noncued repetitive training. We selected reaching movements involving flexion–extension patterns about the elbow in a horizontal plane because they are core components of many functional reaching movements of daily activities. Secondly, we applied a mathematical optimization model to identify potential mechanisms of change, induced by rhythm, in the brain's control strategy of the arm movements. The comprehensive dynamic changes in spatiotemporal and force parameters during rhythmic gait training strongly suggest that a simple trigger or pacing function can only insufficiently explain the effect of rhythm on motor control [29]. Movements patterns are generally considered to be planned around optimization principles in the central nervous system [48,49]. Optimization modeling of control factors in movement, related to minimizing certain physiological or kinematic cost functions, has been successfully used by computational brain theory to understand motor control from a dynamic systems control view [5], in which the brain tries to build optimal strategies for task-oriented problem solving. From basic physiological research, it is known that (a) auditory input can raise spinal motor neuron excitability to increase motor readiness before supraspinal input occurs [27,33], and (b) auditory rhythm rapidly creates stable perceptual traces as anticipatory time schema which attract and rapidly entrain the periodicity of motor patterns [33,42,44]. Periodicity entrainment in rhythmic cuing is based on direct coupling of the motor response to the sensory input, similar to entrainment in coupled oscillator models. Periodicity entraiment further implies that the brain, in addition to aligning the movement endpoints in phase with the stimulus signal, uses period adaptation, i.e. the interval or duration function of the rhythm, to synchronize the execution time of the movement to the stimulus interval. This means that enhanced temporal information is provided by the rhythmic cue to the brain as a continuous time reference at any stage of the movement throughout the entire rhythmic beat cycle [42]. Based on previous research, we studied arm acceleration profiles in a mathematical optimization model to show if enhanced and stable temporal information flow from the rhythm allows the brain to map and scale smoother temporal parameters of positional change across the entire movement cycle of the paretic arm, resulting in the spatiotemporal regulation of the entire movement trajectory.

2. Methodology

2.1. Subjects

A total of 21 right-handed patients (eight women, 13 men, mean age 52.7 ± 13.7 years) participated in the study after giving informed consent. All patients had suffered from MR-confirmed left hemispheric CVAs 4–19 months (mean 11.4 ± 52 months) prior to admission to the study and were attending out-patient therapy programs. A total of 19 patients had suffered ischemic strokes, with 15 in the middle cerebral artery distribution and four in the anterior cerebral artery distribution. Two patients had suffered an intracerebral hemorrhage related to a cerebral aneurysm (anterior communicating artery). Careful neurological and neuropsychological evaluation prior to the study had ruled out the presence of hemi-neglect, attentional, speech, or sensory deficits in all patients at this stage of their recovery. Deficits in general fine motor control of the hand and discrete functional use of the digits in hand dexterity, however, were present in all patients. Hemiparetic upper limb spasticity was present in mild form in nine patients (stages 4–5 of the Brunnstrom-Scale of Spasticity and Motor Recovery), and in moderate form in 12 patients (stage 3–4), affecting mostly elbow and wrist flexion/extension and shoulder displacement. Baseline elbow and shoulder ranges of motion are summarized in Section 5.

3. Experimental procedure

Patients were seated comfortably in a chair approximately 30 cm in front of a height-adjustable table on whose surface two touch-sensitive sensor switches were mounted, vertically aligned in a saggital plane through the shoulder joint of the paretic side. The sensor spacing was adjusted for each patient as the approximate distance between a patient's elbow point (olecranon of the ulna) and maximal extension of the tip of the middle finger (third digit). Patients were asked to move their arm back and forth for 30 s as evenly timed as possible between the sensors, briefly touching the sensor surface (5 cm × 5 cm) before reversing the arm motion. Patients completed a trial with auditory rhythmic cuing and a trial without any external time cuing.

To control for order effects in performance rhythmic and noncued trials were counterbalanced between patients, i.e. half the patients performed the rhythmic trial first and the noncued trial second with the other half of the patients performing the trials in reverse order. Trial order was randomly assigned between patients. The frequency of the rhythmic cue was matched to the patient's self-paced movement frequency which was assessed before the start of the trial. The auditory rhythm consisted of a metronome-like 1000 Hz square wave tone with a 50 ms plateau time produced by a computerized MIDI-sequencing sound software (Logic 2.5). During rhythmic trials, patients were asked to move their arm in time with the rhythm by touching the sensors on the beat. Patients typically started their movements in the rhythmic trials after they had heard the metronome beat two to three times.

4. Data recording

Movement durations were recorded from the voltage coded sensor touch signal. Release and hold time of the sensors was computed separately. All sensor data were recorded on-line on a PC and stored for further data analysis. Time data were recorded at a sampling rate of 1000 Hz.

Arm kinematics were recorded with a three-dimensional (3D) camera based motion analysis system (SELSPOT) using optoelectronic data collection. Four active markers, consisting of infrared light emitting diodes were placed on the shoulder, elbow, wrist, and metacarpophalangeal joint of the paretic arm. Marker recording was carried out by two cameras at a sampling rate of 100 Hz. Cameras were placed at 45° angles diagonally ahead of and behind the patient on a parallel sagittal plane approximately 1.5 m away from the patient. Movement coordinates were calculated with the *x*-axis in the anterioposterior, the *y*-axis in the vertical, and the *z*-axis in the mediolateral direction. Data were analyzed statistically using analysis of variance procedures, dependent sample *t*-tests and nonparametric dependent sample tests (Wilcoxon Signed Rank).

5. Results

5.1. Arm timing

No significant interaction effects or main effects for trial order were found in the data analysis. Mean movement duration time was 1425 ± 185 ms (0.71 Hz) for rhythmic cuing and 1446 ± 289 ms (0.69 Hz) for self-paced cuing. This difference was statistically nonsignificant ($P > 0.05$). Since the metronome frequency was matched to the subject's internal movement frequency, both movement times were expected to be similar. The coefficient of variation (CV: standard deviation/mean duration × 100) for rhythmic cuing was 13 and 20% for no rhythm, indicating a statistically significant reduction in variability of movement timing by 35% (dependent sample *t*-test: $t = 3.205$, $P = 0.013$; nonparametric Wilcoxon Signed Rank test: $z = 0.243$, $P = 0.015$). The source for the variability reduction was predominantly located in the airborne travel time. The CVs for sensor contact were similar for both conditions (17% for rhythm; 16% for no rhythm), but during airborne arm travel rhythmic cuing showed a CV of 14% and no rhythm 22%. The amount of total movement time spent in sensor contact was 19% for rhythm and 17% for no rhythm, indicating that both conditions produced very similar time divisions between stationary sensor touch and motion between sensors.

Temporal loop sums were computed as the cumulative difference between a single movement interval and all other movement intervals to assess temporal variability across consecutive movement repetitions (Fig. 1). Mathematically, the temporal loop sum for trace N (in the sequence of M movement traces) is computed as

$$\mathrm{LS}_N = \sum_{i=1}^{M} |t_N - t_i| \qquad (1)$$

where t_i is the travel time for movement trace i. By mathematical definition, loop sums are dynamic indicators of movement stability, as their decrease or increase continually adjusts to changes in variability in the movement sequence. A decreasing loop sum over consecutive movement traces indicates a continual increase in temporal movement stability or decrease in temporal variability. Time loop sums were

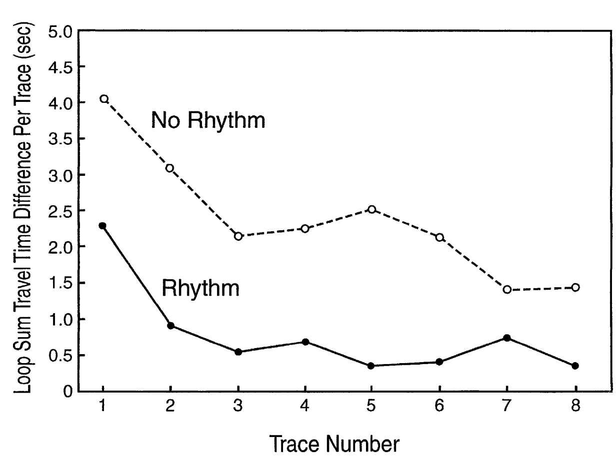


Fig. 1. Group mean loop sums for travel time variability of the paretic arm (*y*-axis) graphed for each movement repetition across trial time (*x*-axis, repetition = trace number). The number of graphed repetitions is adjusted to the slowest moving subject.

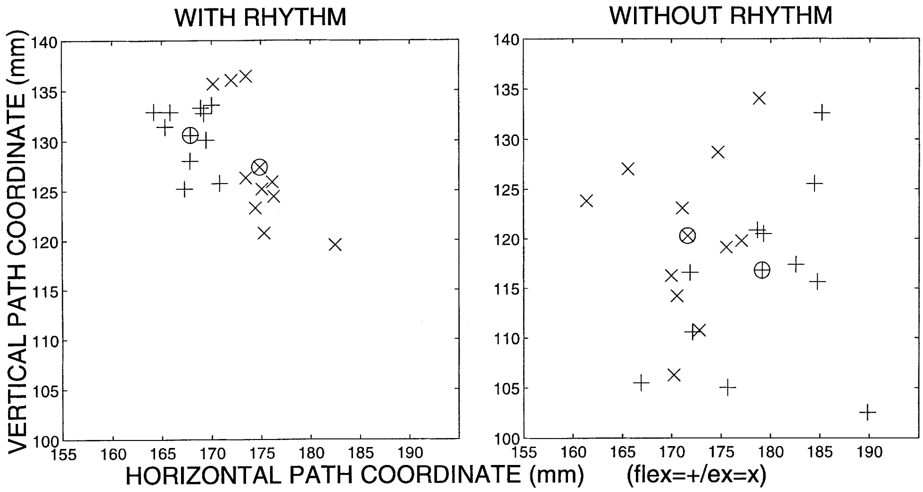


Fig. 2. Scatter plot of a representative trajectory pattern of the wrist of the paretic arm of a stroke patient during reaching in the *yz*-plane at the mid arc cross-section. The *y*-dimension (vertical) is displayed on the *x*-axis, *z* (lateral) on the *y*-axis. The circled dots represent the mathematical midpoint of the trajectory pattern (+: flexion; ×: extension). Clearly visible is the closer clustering of scatter points during rhythm indicating a more consistent arm path with a tendency for segregated clustering for flexion and extension paths not evident during the no-rhythm condition.

already reduced at the outset of the rhythmic trials compared to the nonrhythmic trials and settled on a consistent level after one movement cycle. Without rhythm, loop sums decreased over several cycles but absolute variability rate remained at a higher level without rhythm. Enhanced temporal stability was thus already entrained by the metronome at the outset of each trial without a gradually adaptive learning process.

6. Wrist trajectories

Wrist trajectory deviations in the *yz*-plane were calculated by averaging position data across a trial within a window of ±5% of the path length at the midpoint (50%) and at 20 and 80% of the total trajectory path. The mean coordinate distance was 7.05±3.27 mm with rhythm and 11.84±7.43 mm without rhythm. Trajectory variability expressed as the mean distance from the trajectory midpoint decreased with rhythmic cuing by 40.5%. Coefficients of variations of trajectory variability decreased statistically significant by 26.1% over no rhythm (dependent sample *t*-test: $t = 2.411$, $P = 0.042$; Wilcoxon Signed Rank test: $z = 2.018$, $P = 0.044$). The scatter plot in Fig. 2 of a typical patient's trajectory pattern in the 2-dimensional (2D) *yz*-plane illustrates the closer clustering of the wrist path in the *yz*-plane around the mathematical midpoint of the trajectory with rhythm compared to no rhythm. The *x*-plane is not displayed because it was fixed by the sensor target positions.

Spatial loop sums were calculated as a measure of variability similar to temporal loop sums by

$$\mathrm{LS}_N = \sum_{i=1}^{M} d(N, i) \qquad (2)$$

where $d(N, i)$ is the distance from point N to point i (Fig. 3). The curves displaying the degree of wrist path deviation show the same pattern as the temporal loop sums, indicating coherent and time-synchronized spatio-temporal improvements in the stability of the arm movement across trials with rhythmic cuing. Movement trajectories were already more stable with rhythm than during no rhythm at the outset of the trial. The improvements in temporal and spatial variability during rhythm were highly correlated ($r = 0.82$; $P = 0.001$). No significant correlation was found during the no-rhythm condition.

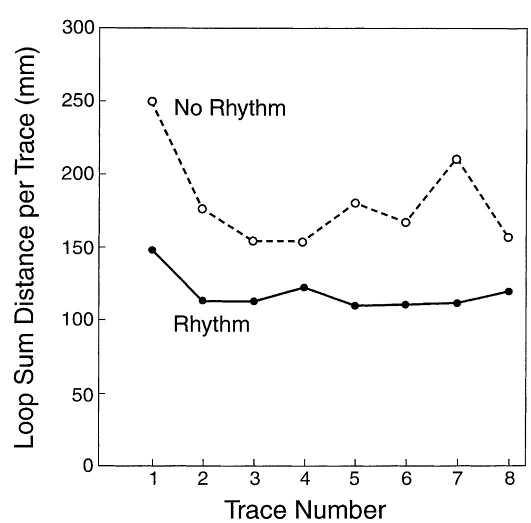


Fig. 3. Group mean loop sums for distance variability of the paretic arm (*y*-axis) graphed for each movement repetition across trial time (*x*-axis, repetition = trace number). The number of graphed repetitions is adjusted to the slowest moving subject.

7. Elbow and shoulder kinematics

Elbow range of motion of the paretic arm showed a statistically significant mean increase in angle by 13.8% (dependent sample *t*-test: $t = 3.44$; $P = 0.007$) between rhythm ($31.47 \pm 9.65°$) and no rhythm ($28.19 \pm 9.96°$). Mean anterior shoulder displacement was 18.7 cm for rhythm and 16.8 cm for no rhythm. This difference was statistically non-significant, indicating that the changes in elbow angles were not due to differential shoulder displacement between the two conditions. The existing displacement in both conditions indicates the presence of some residual synergistic movement patterns in the study sample.

8. Wrist velocity profiles

A total of 11 representative mean profiles of the magnitude of the 3D wrist velocity are shown in Fig. 4. The displayed velocity curves were filtered with a seven-point running average window. Visual analysis revealed that in 16 out of 21 patients, a clear smoothing of the velocity curve occurred during rhythmic cuing by reducing or eliminating the number of reversal peaks in the curve. This was especially prominent during the deceleration phase of the movement. In these patients, a bell-shaped pattern emerged typical for a change in velocity control from an iterative correction model to a kinematically more stable single-pulse model [48]. The five remaining patients showed no change in their profiles.

9. Optimization model of peak acceleration

Acceleration, velocity, and position profiles of the *x*-coordinate of the wrist during a typical movement cycle of a representative subject are displayed in Fig. 5. The movement trajectory depicted is one complete cycle beginning and ending with the arm at full extension touching the rear sensor. The smoother velocity changes are reflected in the acceleration graphs. The dashed line represents the optimized acceleration based on the criterion of minimization of peak absolute acceleration [26,48]. This criterion constrains the acceleration to be constant in magnitude, with a sign change when crossing the midpoint of the motion during both extension and flexion. We applied this model to the acceleration data in our study sample. When computing the deviation of acceleration curves from the optimal (derived under the criterion that the peak absolute acceleration be minimal), the rhythm condition yielded a mean deviation of $38.7 \pm 8.5\%$, while the no-rhythm condition yielded $168.2 \pm 36.1\%$ (dependent sample *t*-test between means: $t = 19.1$; $P < 0.001$), indicating a significantly better model fit for the experimental data cued by rhythm (see Appendices A and B for computation).

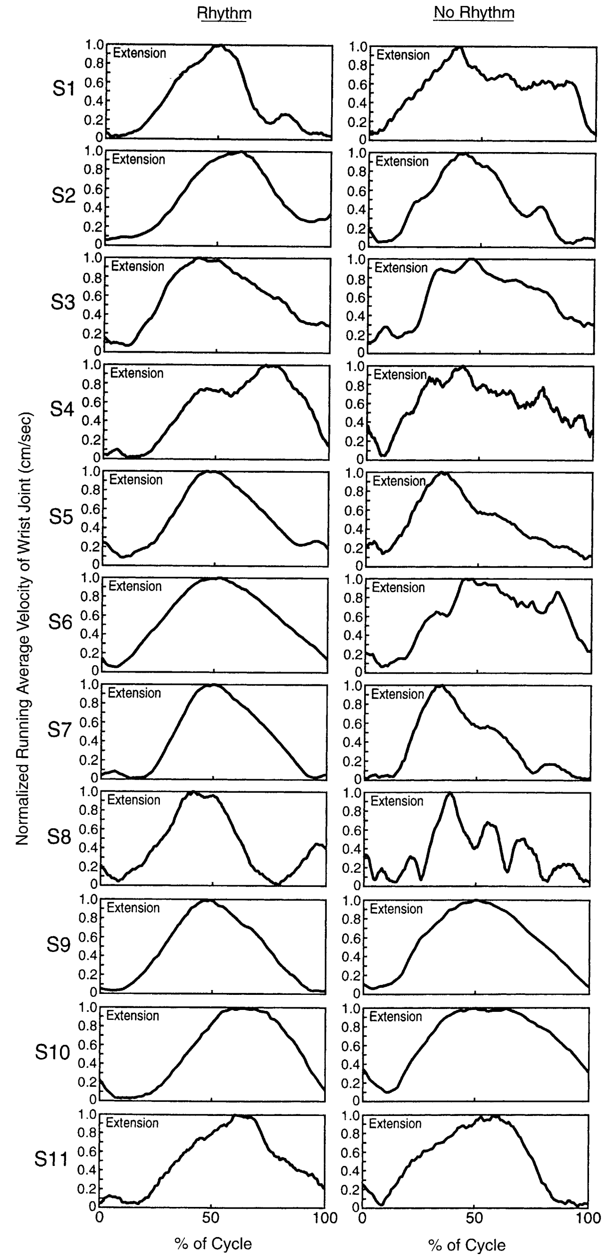


Fig. 4. Representative profiles of 3D velocity magnitudes of the wrist joints of 11 study subjects during the extension phase of the reaching movements between the two target sensors are displayed. The *x*-axis displays 100% of the movement cycle: at 0 and 100% are the sensor surface contacts and reversal points of the movement. The *y*-axis displays the normalized seven-point running average velocity. Note that the summated *xyz*-velocity magnitude of the wrist joint is displayed which does not change the value sign or reach zero velocity at the reversal points. The left column shows the rhythm condition, the right column the no-rhythm condition. Subjects 1–8 exhibit pronounced change in velocity profile, whereas subjects 9–11 do not show change.

10. Rhythmic synchronization

Phase synchronization errors between the onset of the metronome tone and the hand contact at the sensor showed a mean value of -81 ± 176 ms. Data from healthy subjects

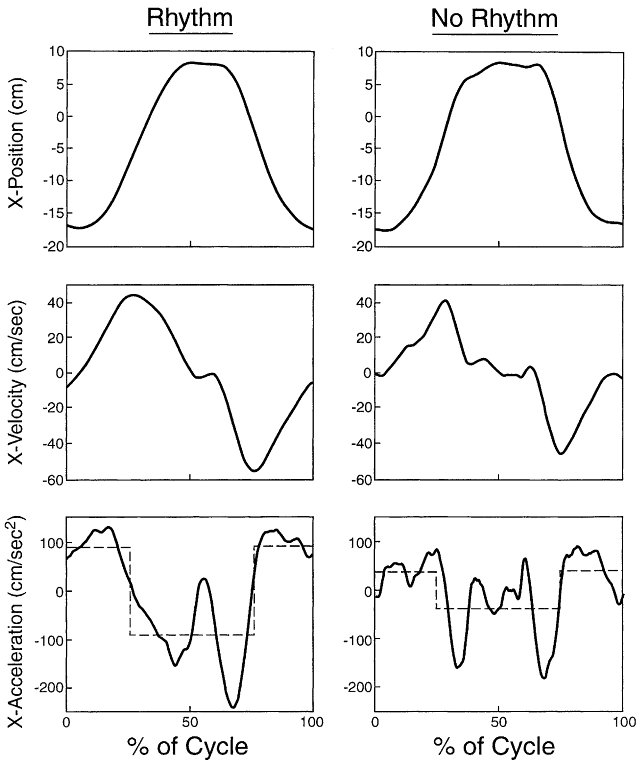


Fig. 5. The *x*-component of the position, velocity, and acceleration curves of the wrist joint are shown for a representative subject's paretic arm during the reaching task. The motion displayed is a full reaching cycle starting and ending at full extension. The *x*-axis is normalized to 100% of the reaching cycle with 0, 50, and 100% marking the sensor surface contacts and reversal points of the motion. The dashed line in the acceleration graph represents the optimized acceleration computed from the formula $a = 16L/T^2$ where L is the distance between targets and T is the cycle period (for derivation see Appendix A; for quantification of curve differences see Appendix B).

show predominantly negative errors as well, typically anticipating the beat onset in a range of 25–40 ms, with variability coefficients at 4–6% of the movement interval [2]. Period synchronization, on the other hand, showed that time deviations between movement intervals and stimulus intervals were rather small for all subjects and averaged to $1.5 \pm 2.1\%$ of the stimulus interval, or 14.4 ± 17.6 ms in real time. These data are in good agreement with data from healthy subjects [42], indicating that stimulus frequencies and movement frequencies were periodically coupled inspite of larger phase offsets. The large mean coefficients of variations in period (13.0%) and phase (12.1%) reflect the existence of considerable fluctuations in time control of the arm. However, the tight periodicity coupling demonstrates that stable rhythmic time information was accessible to the patients during each movement in order to compensate for deviations and maintain essential frequency synchronization.

11. Discussion

Rhythmic auditory-motor synchronization has been successfully modeled in coupled, limit-cycle oscillator models [16,21,32,42,43,47]. These models require for successful motor entrainment that the movement in question is organized in a rhythmic pattern, and that the frequency of the rhythmic timekeeper is, at least until synchronization is achieved, resonant with the natural frequency of the entrained motor response in order to produce the greatest kinematic stability [16,25]. Empirical support for these requirements comes, e.g. from mismatched clock frequency studies with stroke patients [45], and studies in which rhythmic arm movements of Parkinsonian patients improved with a metronome cue [12], whereas nonrhythmic reaching motions to a target with maximal speed and accuracy, using auditory cues as 'go' and 'stop' signals, did not improve motor learning [29].

Our data show support for these models in rehabilitative motor training with stroke patients. The observed changes in timing and trajectory control strongly suggest that the structured time information in auditory rhythm added significant kinematic stability to the patient's paretic arm reaching motions. These changes were not present during the non-rhythmic condition. Flexion–extension patterns are essential components of a host of functional arm movements during activities of daily living which therapy aims to restore in the stroke patients. Our data suggest, therefore, that auditory rhythm may offer an essential component of enhanced sensorimotor control to make hemiparetic arm training more effective. This evidence is further supported by the increased elbow extension angles during the rhythmic condition. Reduction in elbow range of motion due to upper limb muscle spasticity is a serious detriment to functional use of the paretic arm. Temporal regulation of arm trajectory control resulted in reduced spastic inhibition, possibly due to the facilitating effect of rhythm on anticipatory motor planning and execution. Reductions in variability of timing and magnitude of EMG during rhythmic facilitation, indicating more consistent and synchronized motor neuron recruitment patterns, have been observed previously in hemiparetic gait rehabilitation [39,41].

In physical rehabilitation of stroke-related motor deficits, it is important to distinguish between spontaneous neurological recovery processes of lesioned neural systems, and the capacity for functional adaptations of motor behavior through training and learning. Recent research evidence shows that improvements in functional motor skills can still occur through training and learning, although neurological recovery has stabilized and physiological deficits persist [9]. There is also some recent evidence that training and learning paradigms in rehabilitation protocols can actually influence functional plasticity in cortical organization after stroke-related brain damage [8]. In motor learning processes, distinctions have been suggested between the early phases of acquisition or motor program selection and the late phases in which consolidation of skill acquisition and the evolution of motor memory takes place [20]. Recent results have shown that stroke patients, especially with damage in the anterior circulation system, can benefit from training protocols involving motor skill learning over time to improve functional motor control [50]. However, research also shows that inspite of the preserved learning capacities patients with stroke consistently perform with sub-optimal motor control and execution routines (i.e. motor programs), indicating deficits in the acquisition phase and execution process of the motor skill [50]. Our current data show significant immediate improvements in stability of spatiotemporal kinematic parameters with rhythmic patterning versus no temporal sensory cues. Therefore, a role for rhythmic temporal cuing may be indicated in the early stages of motor acquisition and execution. By facilitating the selection and formation of more optimal motor routines during rehabilitative training, the preserved learning process of the patient will be used more effectively. The crucial role of correct and stable temporal organization and rhythmicity in the early stages of motor skill acquisition to optimize learning strategies is already widely recognized in high-performance skill training in music and sports where rhythmic auditory-motor transformations are used to build predictable temporal structures for learning of complex motor programs [11,35,46]. Thus, rhythmic facilitation may make specific contributions to the distinct phases of skill acquisition and consolidation during motor learning by enhancing optimal motor program formation.

In rhythmic response analysis, the close period synchronization within normal variability parameters demonstrates that hemispheric stroke patients can successfully utilize rhythmic stimuli for temporally-based control strategies in rehabilitative motor learning. The larger movement-by-movement fluctuations in the phase response may have been in part due to a compromised final common pathway for motor execution in stroke patients. However, the rhythmic frequency entrainment, associated with central time perception processes, led to significantly reduced spatiotemporal instabilities in hemiparetic arm control.

From a theoretical mathematical–biophysical view point the requirement of synonymity between stimulus interval and movement duration (frequency entrainment) imposes a temporal boundary condition on the physical and neurological problem of moving the arm between two targets as given in our experiment. Once the time constraint is added through the rhythm, the brain is presented with a well-defined optimization problem: how to move the arm from target A to B while minimizing some objective function related to the body's cost in making such a movement [5]. Such cost factors have included total energy output, total muscle force, total power output, peak acceleration, squared joint torque, 'smoothness' of movement and jerk cost, among others [26,48]. We have applied the minimization of peak absolute acceleration over the entire movement cycle as a frequently used optimization criterion to quantify our comparative analysis of wrist kinematics (Fig. 5). Such minimization of peak acceleration results in a unique acceleration–time curve over the movement cycle. Once the acceleration history is known and combined with the temporal (in our study: rhythmic) and spatial (target position) boundary conditions, both the velocity and position–time curves of the arm are known. Our data show that by using an optimization model with a given criterion, rhythmic sensory timing can act as an external forcing function to determine the optimal path trajectory in the absence of sufficient internal control mechanisms. Previous experimental evidence of rhythm's effect on gait kinematics [30] gives further empirical support to this model.

In summary, this study provides significant new evidence that a rhythmic-temporal model of rehabilitative motor training which has previously been utilized in gait rehabilitation can be effectively adapted to improve functional movements of the paretic arm in stroke patients. Our mathematical model suggests that the improvement in spatiotemporal movement kinematics is based on enhanced patterned temporal information flow from rhythmic auditory networks to motor control networks in the brain [7,37,38]. The results are notable in comparison to other rehabilitation techniques because the rhythmic facilitation effect on movement is not mediated by changes in the physical or mechanical constraints of the environment. Previously, rhythmic entrainment has shown significant functional improvements in hemiparetic gait training with stroke patients [17,39] as well as with other movement disorders [40]. These benefits were also significantly higher when compared to other gait training techniques. The immediate benefit of rhythmic cuing on hemiparetic arm control—in conjunction with previous evidence from gait research—provides a strong rationale to apply rhythmic entrainment to the recovery of arm function in long-term hemiparetic stroke rehabilitation. However, further studies, using randomized clinical trials, are necessary to assess the long-term benefits of rhythmic facilitation in hemiparetic arm training in comparison to or in conjunction with other forms of motor therapy.

Acknowledgements

This research was supported in part by a grant from the Deutsche Forschungsgesellschaft, Sonderforschungsbereich 194 to Thaut and Hoemberg (DFG: German Research Council, Special Research Section 194).

Appendix A

The formula for the optimized acceleration a is derived below. In this derivation, one movement cycle of period T is being considered. The movement cycle consists of touching the near target at time $t = 0$, touching the far target at time $t = T/2$, and again touching the near target at time $t = T$. This corresponds physically to extension and flexion of the elbow joint. The position variable under consideration is the location of the middle fingertip along a coordinate axis extending through the centers of the two targets.

The optimization criterion, minimization of the maximum value (peak value) of the absolute value of the acceleration, places constraints on the geometrical configuration of the acceleration–time curve: the acceleration has constant values of $+a$ from $t = 0$ to $t = T/4$, $-a$ from $t = T/4$ to $t = 3T/4$, and $+a$ from $t = 3T/4$ to $t = T$.

Once the acceleration history is established, conclusions regarding the velocity history may be drawn. The velocity has value $v = 0$ at $t = 0$ (contact with near target) and at $t = T/2$ (contact with far target), and reaches a maximum value of $v = v_{\max}$ at $t = T/4$. Furthermore, the velocity varies linearly with time between the points, because constant acceleration implies constant slope for the velocity–time curve. The value of $v_{\max}$ is equal to the (rectangular) area under the acceleration–time curve between $t = 0$ and $t = T/4$

$$v_{\max} = a\frac{T}{4} \tag{1}$$

The distance between the two targets is equal to L, so at $t = T/2$, $x = L$. Now knowing the velocity history, the position history may be determined. The position at $t = T/2$ is numerically equal to the area under the velocity–time curve between $t = 0$ and $t = T/2$. This area may be divided into two congruent triangles with base $T/4$ and altitude $v_{\max}$. Thus,

$$L = 2\frac{1}{2}\frac{T}{4}v_{\max} = \frac{1}{2}\frac{T}{4}\frac{aT}{4} = \frac{aT^2}{16} \tag{2}$$

This may be solved for a to yield

$$a = \frac{16L}{T^2} \tag{3}$$

Appendix B

The extent by which the acceleration deviates from the optimal (derived under the criteria that the peak absolute acceleration be minimal) was quantified by the following process.

1. For each data point, compute the absolute value of the difference between the actual acceleration value and the optimal acceleration. Sum over all the data points for each trial.
2. Divide this sum by the magnitude of the optimal acceleration. This will create a 'percent' deviation which takes into account differences in optimal values.
3. Normalize for the size of the data set by dividing by the number of data points. The result is then an average deviation per data point from the optimal acceleration.

References

[1] Asanuma H, Keller A. Neurobiological basis of motor learning and memory. Concurrent Neuroscience 1991;2:1–30.

[2] Aschersleben G, Prinz W. Synchronizing actions with events: the role of sensory information. Perception and Psychophysics 1995;57:305–17.

[3] Bach-y-Rita P. Recovery from brain damage. Journal of Neurologic Rehabilitation 1992;6:101–200.

[4] Bonita R, Beaglehole R. Recovery of motor function after stroke. Stroke 1988;19:1497–500.

[5] Bryson AE, Ho YC. Applied optimal control. New York: Wiley, 1975.

[6] Buetefish C, Hummelsheim H, Denzler P, Mauritz KH. Repetitive training of isolated movements improves the outcome of motor rehabilitation of the centrally paretic hand. Journal Neurological Sciences 1995;130:59–68.

[7] Buonomano DV, Merzenich MM. Temporal information processing transformed into a spatial code by a neural network with realistic properties. Science 1995;267:1028–30.

[8] Chollet F, DiPiero V, Wise RJS, Brooks DJ, Dolan RJ, Frackowiak RSJ. The functional anatomy of motor recovery after stroke in humans: a study with positron emission tomography. Annals of Neurology 1991;29:63–71.

[9] Dam M, Tonin P, Casson S, Ermani M, Pizzolato G, Iaia V, et al. The effects of long-term rehabilitation therapy on poststroke hemiplegic patients. Stroke 1993;24:1186–91.

[10] Dobkin BH. The rehabilitation of the elderly stroke patient. Clinical Geriatric Medicine 1991;7:507–23.

[11] Effenberg AO. Sonification—Ein komplettes akustisches Informationskonzept fuer die Motorikforschung. In: Krug J, Minnow HJ, editors. Sportliche Leistung und Training. Leipzig. Germany: Academia Publishers, 1995. p. 169–74.

[12] Georgiou N, Iansek R, Bradshaw JL, Phillips JG, Mattingly JB, Bradshaw JA. An evaluation of the role of interbal cues in the pathogenesis of Parkinsonian hypokinesia. Brain 1993;116:1578–87.

[13] Good DC. Treatment strategies for enhancing motor recovery in stroke rehabilitation. Journal of Neurologic Rehabilitation 1994;8:177–86.

[14] Grillner S. Neurological bases of rhythmic motor acts in vertebrates. Science 1985;228:143–9.

[15] Hauptmann B, Hummelsheim H. Facilitation of motor evoked potentials in hand extensor muscles of stroke patients: correlation to the level of voluntary contraction. Electroencephalography and Clinical Neurophysiology 1996;101:387–94.

[16] Hoppensteadt FC, Izhikevich EM. Synaptic organizations and dynamical properties of weakly connected neural oscillators: analysis of a canonical model. Biological Cybernetics 1996;75:117–27.

[17] Hummelsheim H. Rationales for improving motor function. Current Opinion in Neurology 1999;12:697–701.

[18] Hummelsheim H, Arnberger S, Mauritz KH. The influence of EMG-initiated electrical muscle stimulation on motor recovery of the centrally paretic hand. European Journal of Neurology 1996;3:245–54.

[19] Jeffery DR, Good DC. Rehabilitation of the stroke patient. Current Opinion in Neurology 1995;8:62–8.

[20] Karni A, Meyer G, Rey-Hipplito C, Jezztrd P, Adams MM, Turner R, et al. The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. Proceedings of the National Academy of Sciences 1988;95:861–8.

[21] L W, Rand RH. 1:1 and 2:1 phase entrainment in a system of two coupled limit cycle oscillators. Journal of Mathematical Biology 1984;20:133–52.

[22] Keller A, Arissian K, Asanuma H. Synaptic proliferation in the motor cortex of adult cats after long-term thalamic stimulation. Journal of Neurophysiology 1992;68:295–308.

[23] Langton Hewer R. Rehabilitation after stroke. Quarterly Journal of Medicine 1990;76:659–74.

[24] Mandel AR, Nymark JR, Balmer SJ, Grinnell DM, O'Riain MD. Electromyographic feedback versus rhythmic positional biofeedback in computerized gait retraining with stroke patients. Archives of Physical Medicine and Rehabilitation 1990;71:649–54.

[25] Mitra S, Riley MA, Turvey MT. Chaos in human rhythmic movement. Journal of Motor Behavior 1997;29:195–8.
[26] Nagasaki H. Asymmetric velocity and acceleration profiles of human arm movements. Experimental Brain Research 1989;74:319–26.
[27] Paltsev YI, Elner AM. Change in the functional state of the segmental apparatus of the spinal cord under the influence of sound stimuli and its role in voluntary movement. Biophysics 1967;12:1219–26.
[28] Pascual-Leone A, Cohen LG, Dang N, Brasil-Neto JP, Cammarota A, Hallett M. Acquisition of fine motor skills in humans is associated with the modulation of cortical motor outputs. Neurology 1993;43(Suppl 2):A157.
[29] Platz T, Brown RG, Marsden CD. Training improves the speed of aimed movements in Parkinson's disease. Brain 1998;121:505–14.
[30] Prassas SG, Thaut MH, McIntosh GC, Rice RR. Effect of auditory rhythmic cuing on gait kinematic parameters in hemiparetic gait of stroke patients. Gait Posture 1997;6:218–23.
[31] Price SJ, Reding MJ. Physical therapy philosophies and strategies. In: Good DC, Couch JR, editors. Handbook of neurorehabilitation. New York: Marcel Dekker, 1990. p. 181–96.
[32] Rand RH, Cohen AH, Holmes PJ. Systems of coupled oscillators as models of central pattern generators. In: Cohen AH, Rossignol S, Grillner S, editors. Neural control of rhythmic movements in vertebrates. New York: Wiley, 1990. p. 333–67.
[33] Rossignol S, Melvill Jones G. Audiospinal influences in man studied by the H-reflex and its possible role in rhythmic movement synchronized to sound. Electroencephalography and Clinical Neurophysiology 1976;41:83–92.
[34] Sakamoto T, Porter LL, Asanuma H. Long lasting potentiation of synaptic potentials in the motor cortex produced by stimulation of the sensory cortex in the cat: a basis of motor learning. Brain Research 1989;413:360–4.
[35] Shaffer LH. Rhythm and timing in skill. Psychological Review 1982;89:109–22.
[36] Smith DL, Akhtar AJ, Garraway WM. Motor function after stroke. Age and Aging 1985;14:46–8.
[37] Stephan KM, Thaut MH, Wunderlich G, Schicks W, Tian B, Tellmann L, et al. Rhythmic auditory tracking: evidence for three distinct cerebellar circuits. Proceedings of the Society for Neuroscience 1999;148:10.
[38] Stephan KM, Thaut MH, Wunderlich G, Schicks W, Tian B, Tellmann L, et al. Different awareness levels of sensorimotor processing involve distinct anatomical areas with prefrontal cortex. Neuroimage 1999;9:401.
[39] Thaut MH, McIntosh GC, Rice RR. Rhythmic facilitation of gait training in hemiparetic stroke rehabilitation. Journal of Neurological Sciences 1997;151:207–12.
[40] Thaut MH, McIntosh GC, Rice RR, Miller RA, Rathbun J, Brault JM. Rhythmic auditory stimulation in gait training with Parkinson's disease patients. Movement Disorders 1996;11:193–200.
[41] Thaut MH, McIntosh GC, Rice RR, Prassas SG. Effect of rhythmic auditory cuing on temporal stride parameters and EMG patterns in hemiparetic gait of stroke patients. Journal of Neurologic Rehabilitation 1993;7:9–16.
[42] Thaut MH, Miller RA, Schauer ML. Multiple synchronization strategies in rhythmic sensorimotor tasks: period versus phase corrections. Biological Cybernetics 1998;79:241–50.
[43] Thaut MH, Schauer ML. Weakly coupled oscillators in rhythmic motor synchronization. Proceedings of the Society for Neuroscience 1997;298:20.
[44] Thaut MH, Tian B, Azimi MR. Rhythmic finger tapping to cosine wave modulated metronome sequences: evidence for subliminal entrainment. Human Movement Science 1998;17:839–63.
[45] Thaut MH, Ueno K, Hurt CP, McIntosh GC, Hoemberg V. Bilateral limb entrainment and rhythmic synchronization during paretic arm movement in stroke patients. Proceedings of the Society for Neuroscience 1999;365:4.
[46] Thiem B, Green D, Thaut MH. Effect of peripheral timing on left arm EMG patterns in cello performance. Medical Problems of Performing Artists 1994;9:89–96.
[47] Turvey MT, Schmidt RC, Rosenblum LD. Clock and motor components in absolute coordination of rhythmic movements. Neuroscience 1989;33:1–10.
[48] Uno Y, Kawato M, Suzuki R. Formation and control of optimal trajectory in human multijoint arm movements. Biological Cybernetics 1989;61:89–101.
[49] Wada Y, Koike Y, Vatikiotis-Bateson E, Kawato M. A computational theory for movement pattern recognition based on optimal movement pattern generation. Biological Cybernetics 1995;73:15–25.
[50] Winstein CJ, Merians AS, Sullivan KJ. Motor learning after unilateral brain damage. Neuropsychologia 1999;37:975–87.

Rhythmic Auditory Stimulation Improves Gait More Than NDT/Bobath Training in Near-Ambulatory Patients Early Poststroke: A Single-Blind, Randomized Trial

M. H. Thaut, PhD, A. K. Leins, PhD, R. R. Rice, MS, H. Argstatter, MA, G. P. Kenyon, MS, G. C. McIntosh, MD, H. V. Bolay, PhD, and M. Fetter, MD

Objectives. The effectiveness of 2 different types of gait training in stroke rehabilitation, rhythmic auditory stimulation (RAS) versus neurodevelopmental therapy (NDT)/Bobath–based training, was compared in 2 groups of hemiparetic stroke patients over a 3-week period of daily training (RAS group, n = 43; NDT/Bobath group =35). *Methods.* Mean entry date into the study was 21.3 days poststroke for the RAS group and 22.3 days for the control group. Patients entered the study as soon as they were able to complete 5 stride cycles with handheld assistance. Patients were closely equated by age, gender, and lesion site. Motor function in both groups was preassessed by the Barthel Index and the Fugl-Meyer Scales. *Results.* Pre- to posttest measures showed a significant improvement in the RAS group for velocity (P = .006), stride length (P = .0001), cadence (P = .0001) and symmetry (P = .0049) over the NDT/Bobath group. Effect sizes for RAS over NDT/Bobath training were 13.1 m/min for velocity, 0.18 m for stride length, and 19 steps/min for cadence. *Conclusions.* The data show that after 3 weeks of gait training, RAS is an effective therapeutic method to enhance gait training in hemiparetic stroke rehabilitation. Gains were significantly higher for RAS compared to NDT/Bobath training.

From Molecular, Cellular, and Integrative Neuroscience Programs, Colorado State University, Fort Collins (MHT); University of Applied Sciences Heidelberg, Heidelberg, Germany (AKL, HA, HVB); Poudre Valley Hospital, Fort Collins, CO (RRR, GCM); Colorado State University, Fort Collins (GPK); SRH-Hospitals Karlsbad-Langensteinbach, Karlsbad, Germany (MF).

Address correspondence to Michael H. Thaut, PhD, Molecular, Cellular, and Integrative Neuroscience Programs, Center for Biomedical Research in Music, 102 Music Building, Colorado State University, Fort Collins, CO 80523. E-mail: michael.thaut@colostate.edu.

Thaut MH, Nickel AK, Rice RR, Argstatter H, Kenyon GP, McIntosh GC, Bolay HV, Fetter M. Rhythmic auditory stimulation improves gait more than NDT/Bobath training in near-ambulatory patients early post stroke: a single-blind, randomized trial. *Neurorehabil Neural Repair* 2007;21:455–459.

DOI: 10.1177/1545968307300523

Key Words: *Gait rehabilitation—Stroke—Rhythmic auditory stimulation.*

Restoration of mobility is critical to successful rehabilitation after stroke, which makes recovery of functional gait a high priority. Each year, 750,000 individuals have a stroke in the United States, a prevalence of 200 to 300 per 100,000 inhabitants.[1,2] The vast majority of individuals in whom at least partial recovery is observed experience persistent problems especially in the area of neurological motor deficits. For example, epidemiological studies have shown that 20% of stroke survivors remain wheelchair bound and 60% show gait deficits of varying degrees.[3] However, intensive rehabilitation programs have shown to improve gait function[4] and continued research into the efficacy of various treatment approaches continues to hold great benefit.

A number of intervention techniques are in current use, based on different models of motor physiology and disease recovery. Traditional but still widely used techniques include neurodevelopmental therapy (NDT), in Europe known as Bobath-therapy, the Brunnstrom method, proprioceptive neuromuscular facilitation (PNF), and the Rood method.[4,5] However, the research evidence base for the effectiveness of one approach over another has not been demonstrated.

One recent form of gait therapy, rhythmic auditory stimulation (RAS), involves the use of rhythmic sensory cuing of the motor system. RAS is based on entrainment models in which rhythmic auditory cues synchronize motor responses into stable time relationships, similar to oscillator coupling models. Rhythm serves as an anticipatory and continuous time reference on which movements are mapped within a stable temporal template. The fast-acting physiological entrainment mechanisms between auditory rhythm and motor response serve as coupling mechanisms to stabilize and regulate gait patterns. In several clinical research studies, RAS significantly improved gait and other movement parameters (eg, upper extremity function) during rehabilitation for hemiparesis.[6-13] RAS can be used as a self-contained training protocol, but its principles of rhythmic cuing and temporal regulation can also be integrated into other interventions.

The purpose of this study was to examine the clinical efficacy of RAS, based on the experimental design of the study by Thaut et al,[7] by comparing 3 weeks of RAS against the NDT/Bobath method. which is one of the most widely used gait therapies.

METHODS

Subject Selection

From an eligible catchment pool of 155 patients, 78 patients from 2 research centers in Germany and the United States were selected by a random number table. Patients were randomly assigned to either the experimental (RAS; n = 43; male = 22, female = 21) or control (neurodevelopmental technique/Bobath; n = 35; male = 19, female = 16) training group (see Table 1). Treatment allocation was accomplished by computerized random number generators in both centers. Random numbers for the allocation-to-treatment sequence were concealed from the recruiter and the therapists carrying out the training. Patients were informed of the 2 possible treatment allocations but blinded to the aims of an experimental versus control condition. Ethical review board clearance was obtained for all patients.

Subject Characteristics

Table 1 describes the patients. Mean age for the RAS group was 69.2 ± 11.5 and for the NDT/Bobath group 69.7 ± 11.2 years. Lesion site was closely matched in both groups. Mild to moderate sensory dysfunction was present in all middle cerebral artery distribution strokes. Both groups had lower limb spasticity, most pronounced in knee flexors/extensors, plantar flexion, and hip flexors/extensors, as typical for a stage 4 or early stage 3 on the Brunnstrom hemiplegia recovery scale.[14]

Subject Assessment and Training

Both groups were assessed by blinded physical therapists who performed the Barthel Index[15] and the Fugl-Meyer Scale.[16] The Fugl-Meyer score was 31.4 for the control group and 33.3 for the RAS group (balance and lower extremity function combined). The Barthel Index score was 45.5 for the control group and 47.5 for the RAS group. Patients entered the study within 4 weeks of onset, as soon as they could complete 5 stride cycles with handheld assistance by the therapist, that is, with no more than support of the forearm, wrist, and elbow at approximately 90 degrees of elbow flexion on the nonparetic side. Handheld assistance was available to all patients throughout training when needed.

Table 1. Subject Characteristics

	RAS	NDT/ Bobath
N	43	35
Age	69.2 ± 11	69.7 ± 11
Gender M/F	22/21	19/16
Side of hemiplegia (R:L)	20:23	16:19
Time between (days) stroke and admission to study	21.3 ± 11	22.2 ± 12
Location of stroke		
Middle cerebral artery	35	30
Internal capsule	4	4
Basal ganglia/thalamus	3	1
Subdural hematoma	1	

RAS = rhythmic auditory stimulation; NDT = neurodevelopmental therapy.

Mean entry date poststroke was 21.3 ± 10.8 days for the RAS group and 22.3 ± 14.7 for the NDT/Bobath group. The study duration was 3 weeks, with gait training daily for 30 minutes, 5 times per week. Four gait therapists for each group conducted the training to ensure consistency in training protocols and procedures. Each center had its own independently trained pool of therapists. Therapists were not blinded to the treatment conditions of the study. However, because both conditions are considered full treatment conditions, no performance bias was expected. Total walking time was tracked in both groups to ensure consistent exercise duration. Pre-gait exercises were not included in the actual training period of the experimental trials and were carried out in similar fashion in both groups if therapeutically indicated.

RAS training followed established protocols[7,17] using a metronome and specifically prepared music tapes in digital MIDI format to ensure temporal precision and tempo stability as well as full capacity for frequency modulation of the stimulus based on patient needs. After an initial cadence assessment, cuing frequencies were matched to the gait cadence for the first quarter of the session. During the second quarter, cue frequencies were increased in 5% increments as

Table 2. Pretest and Posttest Means and Standard Deviations, Mean Differences Within and Between Groups, and 95% Confidence Intervals Around Mean Differences Between Groups

	Pretest Week 0		Posttest Week 3		Differences Within Groups Week 3 – Week 0		Differences Between Groups* Week 6 – Week 0
	Exp	Ctrl	Exp	Ctrl	Exp	Ctrl	
Velocity (m/min)	14.1 (6.3)	13.0 (5.9)	34.5 (9.1)	20.3 (6.5)	20.4	7.3	13.1 (6.9, 19.3)
Stride length (m)	0.53 (0.12)	0.50 (0.12)	0.88 (0.21)	0.67 (0.24)	0.35	0.17	0.18 (0.13, 0.23)
Cadence (steps/min)	53 (10.8)	50 (9.9)	82 (12.9)	60 (9.9)	29 10	19	(10.4, 27.6)
Symmetry (swing ratios)	0.42 (0.12)	0.40 (0.12)	0.58 (0.05)	0.46 (0.07)	0.16	0.06	0.10 (–0.04, 0.24)

*CI boundaries in parentheses.

kinematically indicated without compromising postural and dynamic stability. During the third quarter, adaptive gait patterns, for example, ramp or step walking, were practiced. The last quarter was spent fading the cues intermittently to train for independent carryover. The control group trained the same amount of time and distance, following NDT and Bobath principles as well as using similar instructions about gait parameters to practice, but without rhythmic auditory cuing.

Testing

All patients were tested 1 day before the training sessions started and 1 day after the last training session. All available participant data after removing dropout participants were analyzed in an intention-to-treat analysis. Testing was carried out without RAS present. For testing, patients walked along a 10 m flat walkway. Two meters on either side were available for acceleration and deceleration without data recording. Gait parameters were recorded at a sampling rate of 500/sec with a computerized foot sensor system consisting of 4 foot contact sensors (heel, first metatarsal, fifth metatarsal, big toe) embedded into shoe inserts. Sensor data were stored online in a portable microprocessor and downloaded after the test walk into a PC with interface hardware and analysis software.

RESULTS

The dropout rate in one center was 23% of initially included patients. There was a 10% dropout rate in the other center. Dropout reasons were due to hospital transfer, early discharge, medical complication, or unspecified personal reasons.

Four major gait parameters critical for improved functional gait were measured and statistically analyzed: velocity, stride length, cadence, swing symmetry (calculated as the ratio between the swing times of 2 consecutive steps using the longer step—ie, paretic vs nonparetic leg—as the denominator). After statistical checks for equivalence of variance (Levene's F test) in each parameter, 2-tailed t test comparisons for independent samples were carried out for pretest differences between the RAS group and the NDT/Bobath group. Pre- and posttest means as well as effect size differences and confidence intervals are given in Table 2. At pretest, there were no significant differences between the 2 groups in each parameter: velocity ($df = 76$, $t = 1.01$, $P = .347$), stride length ($df = 76$, $t = 1.75$, $P = .111$), cadence ($df = 76$, $t = 1.49$, $P = .141$), and swing symmetry ($df = 76$, $t = 1.13$, $P = .285$).

After 3 weeks of gait training, t test comparisons for posttest differences between groups were carried out. Significant differences were found in favor of RAS training in all 4 gait parameters: velocity ($df = 76$, $t = 2.83$, $P = .006$), cadence ($df = 76$, $t = 5.13$, $P = .0001$), stride length ($df = 76$, $t = 4.6$, $P = .0001$), and symmetry ($df = 76$, $t = 2.13$, $P = .049$). Effect size analysis showed improvements for RAS over NDT/Bobath training of 13.1 m/min for velocity, 0.18 m in stride length, 19 steps/min in cadence, and 0.10 in gait symmetry (swing ratio) (Table 2).

Data of patient satisfaction showed a significant main effect in favor of the RAS group ($df = 1,24$, $F = 6.35$, $P = .019$). However, both groups showed continued increases in satisfaction ratings across therapy (RAS: 77%-84%-87%; NDT/Bobath: 64%-70%-75%).

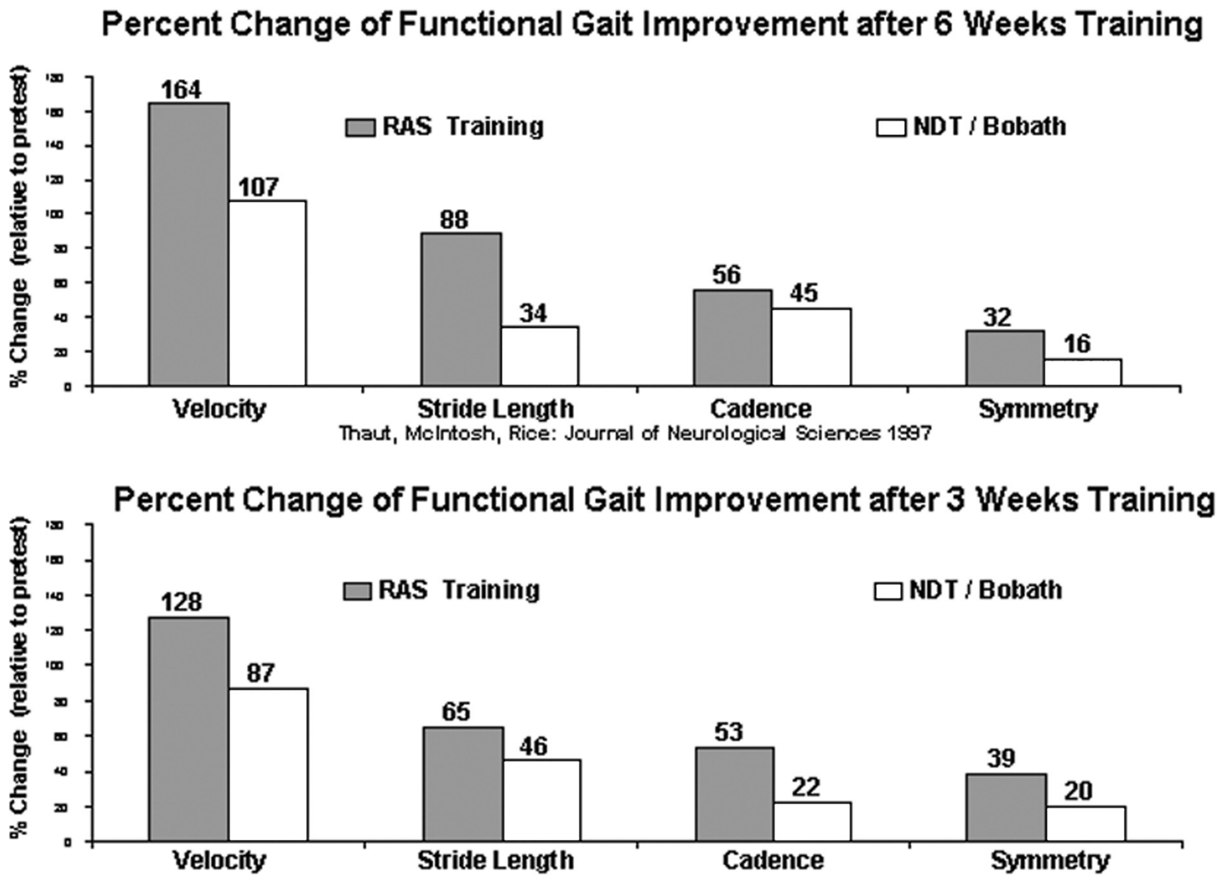


Figure 1. Comparison data for treatment duration.

DISCUSSION

Statistical analysis of 4 gait parameters after a 3-week gait training period in subacute hemiparetic stroke rehabilitation showed significantly greater improvements for training with rhythmic auditory stimulation relative to training within a standard NDT/Bobath protocol. Differences between pre- and posttest as expressed in percent change showed substantial differences between the 2 groups in favor of RAS: velocity 128.8% (RAS) versus 87.6% (NDT/Bobath); stride length 65.9% versus 46.1%; cadence 53.8% versus 22.2%; symmetry 39.1% versus 20.2% (Figure 1). Noteworthy is that RAS training produced mean effects substantially higher in improvement over NDT/Bobath training: 13.1 m/min for velocity, 0.18 m for stride length, 19 steps/min for cadence, and 0.10 in symmetry (swing ratio) over NDT/Bobath training (Table 2).

In the RAS group, significant increases in gait velocity were driven by somewhat larger increases in stride length than cadence. In the control group, a similar pattern of contribution was observed; however, with a much larger differential in magnitude for stride length relative to cadence, which only improved by 22% on average. Because changes in stride length and cadence are kinematically linked in healthy gait, increases in those parameters that are coupled more closely may suggest a more functional recovery of gait mechanics. Improvements in velocity that are mostly driven by stride length and a disproportionally smaller cadence change may indicate uncoupling of kinematic linkages due to compromised asymmetric step patterns.[4]

Although substantial increases in swing symmetry were seen for RAS relative to NDT/Bobath, the smaller improvement compared to the other parameters shows the higher resistance of this parameter to rehabilitation efforts.[1] However, the isochronous nature of the rhythmic timing cue still showed higher efficacy in symmetry restoration than NDT/Bobath alone. Similarly, moderate results in regard to swing symmetry were obtained by our group in a previous study,[7] but improvements in this parameter did not reach statistical significance.

Improvements in velocity, stride length, and cadence were statistically similar to previous data,[7] but with smaller percentage increase rates in this study compared to previous data. Velocity improved 128% versus 164%

in previous research (control group improvement 87.6% vs 107%). The main difference in treatment dosage between the 2 studies was the duration (3 vs 6 weeks). Therefore, the difference in percent improvement suggests that the additional 3 weeks of training had a substantial effect on speed of walking, which is a critical parameter in functional gait recovery. Although the current study was not designed to statistically compare different treatment durations, the similarity in treatment design and diagnostic patient selection criteria allows for a descriptive comparison between the current data and the previous study data,[7] showing the dosage benefit of 6 weeks of therapy over 3 weeks (Figure 1).

When referenced to percent of healthy normative data,[18] results showed substantial differences in favor of RAS. RAS-training patients reached 43.6% of healthy control velocity but only 26.3% in NDT/Bobath, in cadence 66.4% versus 48%, and in stride length 60.5 versus 51.8%.

Considering that predictive states in motor planning, as well as attentional and executive brain networks, reduce performance variability, the intrinsic time structure of rhythmic cues and their almost instantaneous synchronization effect on motor responses can play a critical role in performance regulation by enhancing temporal predictability via interval scaling. It has been shown in optimization models that a rhythmic cue as a predictive time constraint can result in the complete specification of the dynamics of the movement over the entire movement cycle, reducing variability, enhancing temporal precision, and facilitating the selection of optimal movement trajectories, velocity, and acceleration parameters. Thus, temporal-rhythmic motor cues do not only cue speed and timing of movement but also regulate comprehensive spatiotemporal and force parameters[19,20] in restoring motor function in brain rehabilitation.[5,21,22]

In summary, RAS significantly improved gait performance in subacute hemiparetic stroke rehabilitation over NDT/Bobath-based training. The 3-week training period showed smaller overall improvements when compared to a 6-week study with an identical therapy protocol, suggesting the functional importance of additional training for the patient's functional locomotor recovery. Future studies may follow 4 directions to further establish the role of RAS in gait rehabilitation: (a) test other treatment dosages for RAS, (b) compare RAS against other current gait-training methods besides NDT/Bobath, (c) investigate the potential to enhance RAS by adding other current gait therapy techniques, (d) study the effect of RAS in long-term outpatient or community-based settings.

REFERENCES

1. Hesse S, Werner C. Post-stroke motor dysfunction and spasticity: novel pharmacological and physical treatment strategies. *CNS Drugs*. 2003;17:1070-1093.
2. Williams GR, Jiang JG, Matchar DB, Samsa GP. Incidence and occurrence of total (first-ever and recurrent) stroke. *Stroke*. 1999;30:2523-2528.
3. Jorgensen HS, Nakayama H, Pedersen PM, Kammersgaard L, Raaschou HO, Olsen TS. Epidemiology of stroke-related disability. *Clin Geriatr Med*. 1999;15:785-799.
4. Mauritz KH. Gait training in hemiplegia. *Eur J Neurol*. 2002;9:23-29.
5. Hummelsheim H. Rationales for improving motor function. *Curr Opin Neurol*. 1999;12:697-701.
6. Thaut MH, McIntosh GC, Prassas SG, Rice RR. The effect of auditory rhythmic cuing on stride and EMG patterns in hemiparetic gait of stroke patients. *J Neurol Rehabil*. 1993;7:9-16.
7. Thaut MH, McIntosh GC, Rice RR. Rhythmic facilitation of gait training in hemiparetic stroke rehabilitation. *J Neurol Sci*. 1997;151:207-212.
8. Thaut MH, Kenyon GP, Hurt CP, McIntosh GC, Hoemberg V. Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients. *Neuropsychol*. 2002;40:1073-1081.
9. Whitall J, McCombe Waller S, Silver KH, Macko RF. Repetitive bilateral arm training with rhythmic auditory cueing improves motor function in chronic hemiparetic stroke. *Stroke*. 2000;31: 2390-2395.
10. Luft AR, McCombe Waller S, Whitall J, et al. Repetitive bilateral arm training and motor cortex activation in chronic stroke: a randomized controlled trial. *JAMA*. 2004;292:1853-1861.
11. Mandel AR, Nymark JR, Balmer SJ, Grinnell DM, O'Riain MD. Electromyographic feedback versus rhythmic positional biofeedback in computerized gait retraining with stroke patients. *Arch Phys Med Rehabil*. 1990;71:649-654.
12. McCombe WS, Whitall J. Hand dominance and side of stroke affect rehabilitation in chronic stroke. *Clin Rehabil*. 2005;19:544-551.
13. Schauer M, Mauritz KH. Musical motor feedback (MMF) in walking hemiparetic stroke patients: randomized trials of gait improvement. *Clin Rehabil*. 2003;17:713-722.
14. Brunnstrom S. *Movement Therapy in Hemiplegia: A Neurophysiological Approach*. Philadelphia, PA: Harper & Row; 1970.
15. Mahoney FI, Barthel DW. Rehabilitation of the hemiplegic patient: a clinical evaluation. *Arch Phys Med Rehabil*. 1954; 35:359-362.
16. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient: a method for evaluation of physical performance. *Scand J Rehabil Med*. 1975;7:13-31.
17. Thaut MH. *Rhythm, Music, and the Brain: Scientific Foundations and Clinical Applications*. New York: Taylor & Francis; 2005.
18. Oeberg T, Karsznia A, Oeberg K. Basic gait parameters; reference data for normal subjects, 10-79 years of age. *J Rehabil Res Dev*. 1993;30:210-223.
19. Kenyon GP, Thaut MH. Rhythmic-driven optimization of motor control. *Recent Res Dev Biomech*. 2003;1:29-47.
20. Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature*. 1998;194:780-784.
21. Hoemberg V. Evidence-based medicine in neurological rehabilitation—a critical review. *Acta Neurochir*. 2005;93: 3-14.
22. Molinari M, Leggio MG, De Martin M, Cerasa A, Thaut M. Neurobiology of rhythmic motor entrainment. *Ann N Y Acad Sci*. 2003;999:313-321.

APHASIOLOGY, 2008, 22 (7–8), 906–920

Psychology Press
Taylor & Francis Group

The effects of rate control treatment on consonant production accuracy in mild apraxia of speech

Shannon C. Mauszycki and Julie L. Wambaugh

VA Salt Lake City Healthcare System, and University of Utah, Salt Lake City, UT, USA

Background: A primary feature of acquired apraxia of speech (AOS) is a slow speech rate associated with lengthened sound segments and intersegment durations (McNeil, Robin, & Schmidt, 1997). This disturbance in speech production timing has been the focus of a limited number of treatment studies designed to manipulate rate and/or rhythm of speech production with speakers with mild AOS.

Aims: The purpose of this investigation was to study the effects of rate control treatment on sound production accuracy and utterance durations of multisyllabic words, phrases, and sentences in a speaker with mild AOS and aphasia.

Methods & Procedures: An individual with mild AOS and aphasia was trained to produce multisyllabic words and phrases using a combination of metronomic rate control and hand tapping. The speaker was trained to produce one syllable per beat of the metronome in conjunction with hand tapping. Feedback was only provided for accuracy of hand tapping and/or syllable production to the beat of the metronome. No feedback was given regarding the accuracy of sound production. Initially, the speaker's rate of production was reduced, but was then systematically increased. A multiple baseline design was used to examine the acquisition, response generalisation, and maintenance effects of treatment.

Outcomes & Results: Findings revealed an increase in sound production accuracy for trained four-syllable words and some improvement in sound production accuracy for treated phrases and untrained four-syllable words. There was only a slight reduction in total utterance duration for treated items versus untreated items. There was a gradual decline in total utterance duration over time on untrained stimulus generalisation items with no consistent improvement on sound production accuracy.

Conclusions: Treatment resulted in an improvement in sound production accuracy in an individual with AOS and aphasia. Positive changes were observed for treated four-syllable words, phrases, and untrained four-syllable words, although treatment did not directly target sound production accuracy (i.e., feedback was not given regarding accuracy of productions). The study represents an initial investigation of the effects of rate control treatment specifically increasing rate of production in a speaker with mild AOS and aphasia. This type of treatment appears to have promise in terms of improving sound production accuracy and warrants further investigation.

Keywords: Apraxia of speech; Treatment.

Address correspondence to: Shannon C. Mauszycki, Aphasia Research Lab, 151-A, Building 2, 500 Foothill Blvd, Salt Lake City, UT 84148, USA. E-mail: Passbrat@aol.com

This research was supported by Rehabilitation Research and Development, Department of Veterans Affairs. This investigation was presented as a poster at the annual Clinical Aphasiology Conference, Scottsdale, AZ, May 2007. Thanks to Joel Longhurst for his assistance with data analysis.

http://www.psypress.com/aphasiology DOI: 10.1080/02687030701800818

Acquired apraxia of speech (AOS) is a neurogenic sensorimotor speech disorder characterised by disturbed articulation and prosody. A primary feature of AOS is a slow speech rate associated with lengthened sound segments and intersegment durations (McNeil, Robin, & Schmidt, 1997). This disturbance in speech production timing has been the focus of a limited number of treatment studies designed to manipulate rate and/or rhythm of speech production with speakers with AOS.

Treatment techniques utilised to control rate and rhythm have included metronomic pacing (Dworkin & Abkarian, 1996; Dworkin, Abkarian, & Johns, 1988), a combination of metronomic pacing and hand tapping (Wambaugh & Martinez, 2000), pacing board (McHenry & Wilson, 1994), and auditory and/or visual feedback (Brendel, Ziegler, & Deger, 2000; Southwood, 1987). These treatment techniques involve slowing the rate of production of speakers with AOS that already exhibit a slow rate of speech. The reduced rate is thought to allow additional time for motor planning/programming as well as processing feedback. It has also been theorised that the use of rate and rhythm controls may serve to focus the speaker's attention either on increasing accuracy of speech production (Dworkin et al., 1988) or away from speech production (Brendel et al., 2000).

Metronome pacing (Dworkin et al., 1988) was utilised with a speaker with chronic AOS targeting nonspeech repetitive tongue movements, alternating motion rates (AMRs), multisyllabic words, and sentences. Treatment focused on sequentially treating behaviours from the least to the most motorically complex while systematically increasing the speed of production with treatment initiated at extremely slow rates. Findings revealed positive changes on all behaviours treated, but without generalisation across behaviours (i.e., treating of the motorically simple behaviour did not generalise to the more complex behaviour). Metronome pacing was also used by Dworkin and Abkarian (1996) to treat an individual with AOS with significantly impaired phonatory control as well as disturbed articulation and prosody. Treatment focused on repeated practice producing isolated vowels, vowel sequences, and alternating vowel and /h/ sequences with a slowed rate of production that was gradually increased over the course of treatment. There were positive treatment effects for all treated behaviours as well some generalisation to more complex behaviours.

In a later study, metronome pacing and hand tapping (Wambaugh & Martinez, 2000) were combined in a treatment with an apraxic speaker. Treatment involved training multisyllabic words at a rate of production 50% slower than the speaker's typical speaking rate. Treatment involved gradually increasing the speaker's rate of production and incorporating a syncopated rhythm. Temporal measures conducted on a portion of the data revealed no changes in overall durations, although changes in relative durations of syllables were noted. Positive treatment effects were found for trained as well as untrained words in terms of articulatory accuracy, although treatment did not involve training for articulatory accuracy and provided no feedback regarding the accuracy of productions.

A pacing board and clinician cueing (McHenry & Wilson, 1994) have also been employed to reduce the speaking rate with an individual with AOS and dysarthria. The speaker was trained to touch each consecutive section of the board for each unit of speech produced. Treatment involved the imitation of functional phrases and multisyllabic words to structured conversation and unstructured tasks. The authors reported that the reduced rate helped to minimise the behaviours associated with

AOS; however there were no data provided regarding changes in speech production accuracy or intelligibility for the speaker.

In a facilitation study, visual feedback (Southwood, 1987) was utilised with two apraxic speakers. They were instructed to utilise a prolonged manner of speech production while reading passages aloud at a controlled rate of presentation via video display. Findings revealed a decline in speech sound errors for both speakers; however there was no generalisation to other behaviours (i.e., discourse task) and the speakers' speech production were described as unnatural.

Visual feedback was also utilised in a treatment study by Brendel and colleagues (2000). A video display was utilised and speakers were instructed to match their productions to computer-generated rhythmic cues, which were controlled for rate and metrical form. The effects of treatment were examined by number of segmental errors and fluency. Findings revealed an increase in fluency for all speakers and the reduction in segmental errors differed across speakers. Specific information concerning rate manipulations was not provided.

It appears that rate/rhythm treatment techniques have potential to positively impact speech production in AOS, but the lack of replication limits the degree to which these treatments can be advocated in terms of clinical application with speakers with AOS. Only two investigations have directly measured the effects of rate/rhythm treatment on sound production accuracy (Southwood, 1987; Wambaugh & Martinez, 2000) with positive findings for both. Additionally, only one study examined the effects of rhythm control treatment on durational measures, although no differences were found between treated and untreated words (Wambaugh & Martinez, 2000). Furthermore, there has been no investigation that has increased the rate of speech production in a speaker with AOS above their habitual speaking rate. Such increases may be warranted in cases of relatively mild AOS.

The purpose of this investigation was to study the effects of rate and rhythm control treatment by systematically increasing the rate of production with an individual with AOS and aphasia. The acquisition, response generalisation, and maintenance effects of treatment were examined in terms of accuracy of sound production and total utterance duration of multisyllabic words, phrases, and sentences.

METHOD

Participant

The participant was a 35-year-old male with chronic AOS and aphasia. He was 18 months post onset of a cerebral vascular accident and exhibited mild right hemiparesis with associated sensory deficits. The participant was a native English speaker who had 18 years of formal education and prior to his stroke was a mortgage broker. His speech production was consistent with the characteristics of AOS described by McNeil et al. (1997; e.g., consistently reduced rate of speech production, distorted sound errors, segregated syllable production, disrupted prosody). In addition, he demonstrated articulatory groping, repeated production attempts, and awareness of errors. The speaker exhibited mild AOS with a slowed speech rate especially with more complex speech tasks (i.e., conversational and narrative discourse) that required a greater vocabulary with longer words (i.e., four or more syllables) comprised of more complex syllable structure (i.e., two and three consonant clusters). Although there were no phonemes absent from his phonetic inventory, the speaker tended to have an increase in sound production errors with specific phonemes being problematic (i.e., voiceless stops, liquids, and glides) during these tasks. His mean speaking rate was 81 words per minute (wpm) on conversational tasks and also 81 wpm on word and phrase repetition tasks prior to treatment. The participant did not exhibit any signs of dysarthria as described by Duffy (2005). The speaker presented with mild anomic aphasia, as evidenced by his performance on the *Western Aphasia Battery* (Kertesz, 1982). See Table 1 for pretreatment assessment.

Experimental stimuli

A total of 90 words, 40 phrases, and 60 sentences were used to measure the effects of treatment on sound production accuracy (percentage of words correct; PWC) and total utterance duration (TUD). Stimuli comprised seven groups:

(1) 30 four-syllable words with primary stress on second syllable (e.g., obtrusively);
(2) 30 four-syllable words with primary stress on third syllable (e.g., tonsillitis);
(3) 40 four- to five-syllable phrases (e.g., Your hair is lovely);
(4) 10 two- to four-syllable words comprising s-blends (e.g., stringency);
(5) 20 five-syllable words (e.g., dermatologist);
(6) 30 sentences containing a four-syllable word with primary stress on the second syllable (e.g., She obtrusively pushed her idea);
(7) 30 sentences containing a four-syllable word with primary stress on the third syllable (e.g., Tammy had tonsillitis last month).

One-half of the words in groups 1, 2, and 3 were designated as treatment items and the remaining half in each group were designated as response generalisation items. The assignment of items to treatment or no-treatment groups was quasi-random—i.e., lists were similarly matched for phonetic complexity and/or number of syllables (phrases). The items in groups 4, 5, 6, and 7 were not divided into treatment and untreated groups, because these stimuli never received treatment.

TABLE 1
Pretreatment assessment results

Assessment tool	*AOS speaker*
Porch Index of Communicative Ability (Porch, 2001)	
Overall Score	13.3
Percentile Score	79.0
Western Aphasia Battery (Kertesz, 1982)	
Aphasia Quotient	94.0
Classification	Anomic
Assessment of Intelligibility of Dysarthric Speech (Yorkston & Beukelman, 1981)	
Word Level Intelligibility	92%
Apraxia Battery for Adults-2 (Dabul, 2000)	
Level of Impairment	Mild AOS
Coloured Progressive Matrices (Raven, Raven, & Court, 1988)	
Total score (36 possible)	33

Experimental design

A single-subject multiple baseline design across behaviours was used to examine treatment effects. PWC and TUD in the experimental stimuli served as the dependent variables. Baseline probes were conducted until TUD was stable (no differences between mean TUD greater than 0.25 seconds for three baseline probes prior to initiating treatment) for the set of stimuli to be treated first. Treatment was applied sequentially to the different sets of stimuli in the following order: four-syllable words with primary stress on the second syllable, four-syllable words with primary stress on the third syllable, and phrases.

Baseline phase. During baseline, TUD and PWC were measured for each of the 190 experimental words, phrases, and sentences in repetition probes. The 190 items were presented in seven groups (i.e., for each set of stimuli), with the order of groups and the order of stimuli within groups randomised. The experimenter produced each stimulus item and asked the speaker to repeat the item as accurately as possible. No feedback was provided regarding the accuracy of production. Five baselines probes were conducted for all of the stimulus groups except for the two sets of sentences. Only two baseline probes were conducted for the two sets of sentences (one set comprised four-syllable words with second syllable stressed and the other set comprised four-syllable words with third syllable stressed).

Treatment phase. Probes identical to those conducted during baseline were completed at the beginning of each session prior to treatment.

Maintenance and follow-up phase. Maintenance of previously trained behaviours was measured during training of subsequent stimuli. Follow-up probes were conducted for all stimuli at 2 and 4 weeks post treatment.

Scoring. For each stimulus item an accurate and acceptable production was determined via broad phonetic transcription. Each consonant segment in a stimulus item was determined to be correct or incorrect. A stimulus item was deemed correct if no consonant segments were in error (i.e., vowels were not judged in terms of accuracy). If a stimulus item had a consonant segment in error it was deemed incorrect. Then the numbers of correct stimulus items (i.e., with no consonant segments in error) were added up and divided by the total number of stimulus items for that list for the measurement of percentage of words correct (PWC). Only the speaker's first complete production of each stimulus item was analysed, although the speaker frequently attempted additional productions due to incorrect production of the target stimulus item.

All productions were transcribed on-line using broad phonetic transcription. All baseline and probe sessions were recorded and these audio recordings were used to verify transcriptions. Each group of word and phrase stimuli in every probe was scored for a PWC. For sentence stimuli, only the target four-syllable word in each sentence was scored for PWC.

Temporal-acoustic measures. The speech data were analysed using Multi-Speech (Kay Elemetrics, 2004). Spectrographic and oscillographic displays were produced and linked to measure TUD, which was measured from the onset of noise burst to the point of a reduction in amplitude to baseline.

Reliability. A total of 10% of the productions were randomly selected for reanalysis of perceptual and acoustic measurements by another investigator. Point-to-point agreement for broad phonetic transcription was calculated at 89%. Reanalysis of TUD revealed 96% of the measurements fell within 10% (plus or minus) of the original measurement.

In order to control for the influence of rate of production (i.e., model) in probes by the experimenter, a mean rate of production (TUD) was established for stimuli items in baseline probes. Then 10% of the experimenter's productions were randomly selected on subsequent probes for analysis. The analysis revealed that 95% of the TUD for the experimenter were within 10% (plus or minus) of the mean rate of production established in baseline probes.

Treatment

The speaker was trained to produce four-syllable words and four- to five-syllable phrases in rhythm with a metronome in conjunction with hand tapping. Treatment involved the use of a digital metronome (audible click plus small flashing light). For each set of stimuli, the average duration produced by the speaker during five baseline probes prior to initiating treatment on that set of stimuli was used to determine the metronome setting for treatment. Initially, the beats per minute (bpm) were set to reflect an increase of word or phrase duration 10% greater than the speaker's average duration (i.e., speaking rate was slowed). Then, the rate of bpm was set to reflect a decrease of word and phrase duration in 10% increments on two occasions to approximate a more natural rate (i.e., faster rate of speech).

The speaker was trained to produce one syllable per beat while tapping his hand in unison with the metronome. Clinician participation on the first level of treatment involved the clinician providing the speaker with four models of the target treatment item (prior to the speaker's participation) and then three productions of the target treatment item were produced in unison by the clinician and participant. Clinician modelling was systematically faded on the second level of treatment. An important aspect of the treatment was that feedback was never provided regarding the accuracy of sound production. Feedback was provided only for accuracy of tapping and/or syllable production to the beat of the metronome rate (see Appendix for an outline of treatment levels/steps). Two presentations of the 15 treatment items for words (20 items for phrases) constituted one treatment trial. The speaker typically completed one to two treatment trials during each treatment session. The speaker was seen twice a week for approximately 30–45 minutes each session.

Treatment was provided by an ASHA-certified speech language pathologist. Treatment was applied at each treatment level until the participant reached at least 90% accuracy in tapping and syllable production to the beat of the metronome in two consecutive treatment sessions or until 10 treatment sessions were completed.

RESULTS

Figures 1 and 2 depict the acquisition and response generalisation findings relative to TUD (top graph) and PWC (lower graph). These data were derived from

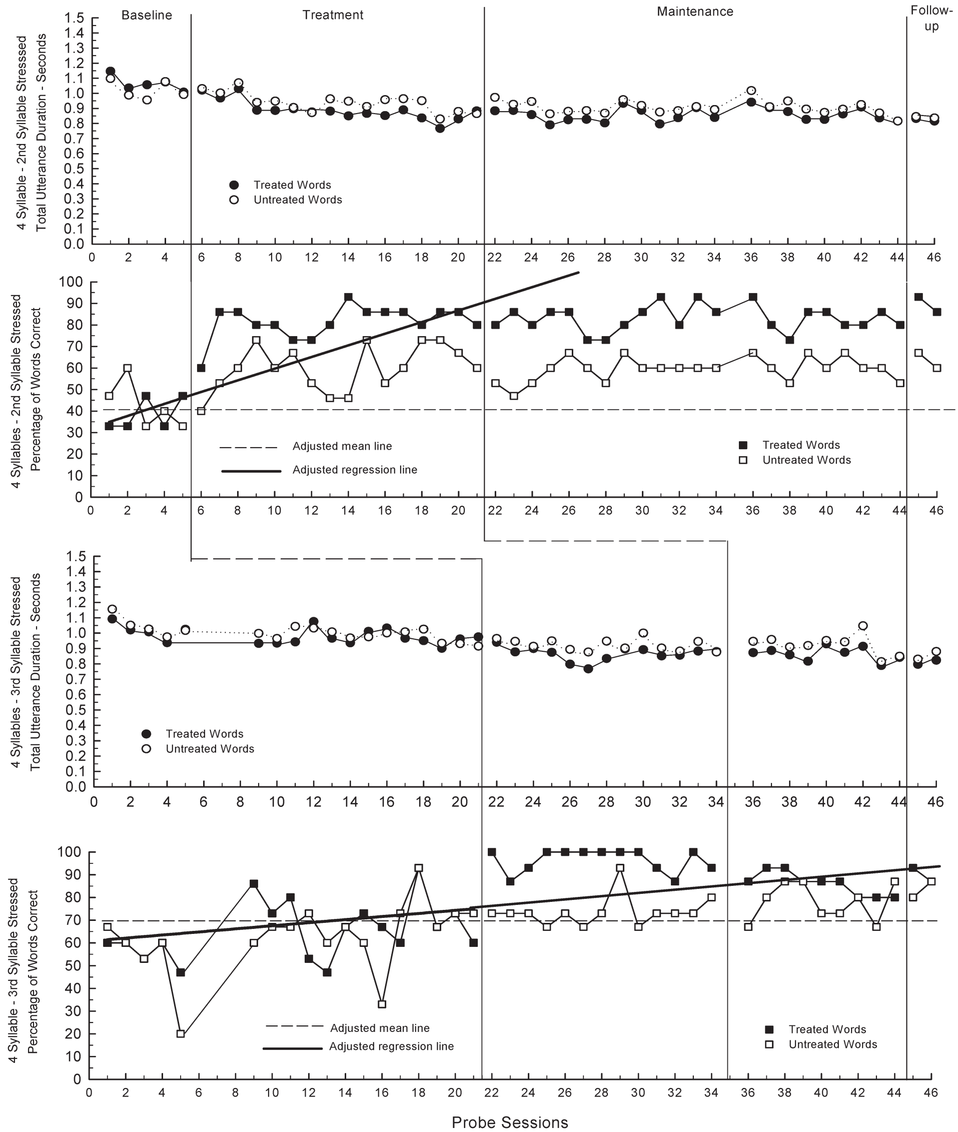


Figure 1. Total utterance duration and percentage of words correct in probes for treated and untreated items on four-syllable words with primary stress on the second syllable (top two graphs) and four-syllable words with primary stress on the third syllable (lower two graphs).

productions during probes (not during treatment). Four-syllable words with primary stress on the second syllable (4 syl.-2nd) are shown in the top two graphs of Figure 1 and four-syllable words with primary stress on the third syllable (4 syl.-3rd) are shown in the lower two graphs of Figure 1, with treated and untreated words displayed separately. For both sets of four-syllable words, TUD was slightly lower for treated words with the application of treatment and was maintained after treatment and post treatment (2 and 4 weeks).

For the 4 syl.-2nd words, PWC averaged 39% during baseline. Following the application of treatment, PWC increased on average to 84% for the final three

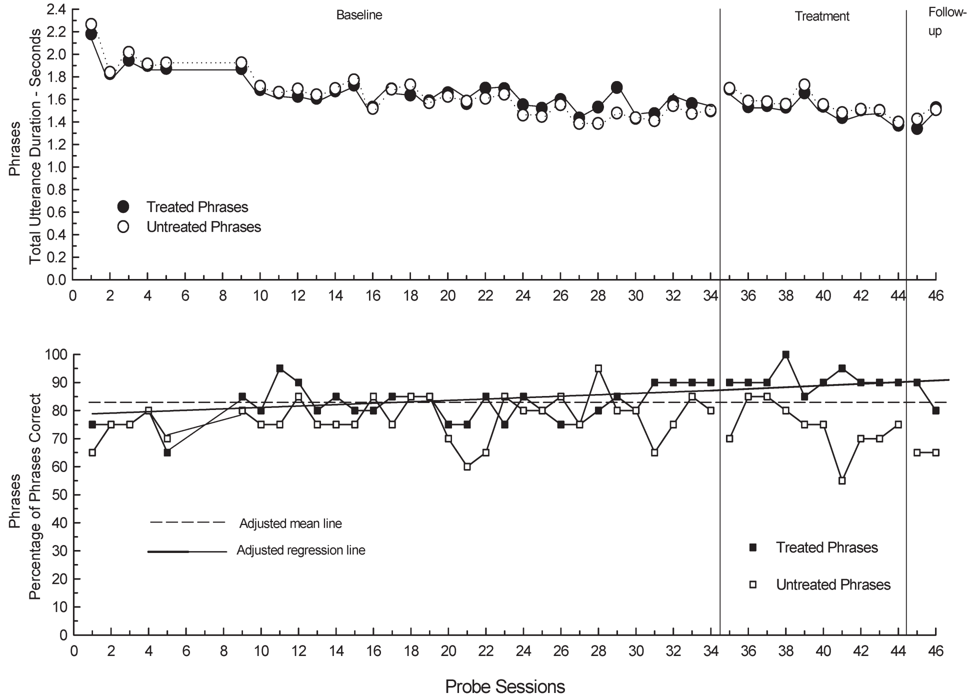


Figure 2. Total utterance duration (top graph) and percentage of phrases correct (lower graph) in probes for treated and untreated phrases.

probes of the treatment phase. PWC was on average 20% greater for treated words in comparison to the untreated words. In order to assist in visual inspection of these data, the conservative dual-criterion method (CDC) described by Fisher, Kelly, and Lomas (2003) was utilised. The CDC has been recommended for use in assisting the interpretation of single case designs, because it has been found to control Type I error better than the split-middle method, the general linear model, and interrupted time series analysis when data are autocorrelated (Bloom, Fischer, & Orme, 2006). The CDC method involves adjusting mean and least square trend lines (note: the lines were adjusted upwards by .25 standard deviations). As shown in Figure 1 (second graph), the lines were extended through the treatment phase. With 16 points in the treatment phase, the CDC method required that at least 12 data points in that phase be above both the criterion lines to conclude that there was a reliable treatment effect. Of the 16 data points during the treatment phase, 12 fell above both criterion lines indicating a positive treatment effect for 4syl.-2nd words. The magnitude of the difference in baseline probe data compared to treatment phase probe data for 4 syl.-2nd words was estimated using the Δ-index calculation which yielded an ES-index of 5.57 (Pearson Education Inc, 2006) representing a large effect size (Beeson & Robey, 2006).

For the 4 syl.-3rd words, PWC averaged 66% during baseline. With the application of treatment, PWC rose to 93% in the final three probes of the treatment phase. PWC was on average 20% greater for treated words in comparison to the untreated words. Again, the CDC method was utilised to assist with visual inspection of this data. In Figure 1 (graph 4), all 13 treatment probe data points fell above both criterion lines, suggesting a reliable treatment effect for 4 syl.-3rd words. The magnitude of difference in baseline probe data compared to treatment phase data was estimated for this stimulus set using the Δ-index calculation, which yielded an ES-index of 2.39 (Pearson Education Inc, 2006) indicating a small effect size (Beeson & Robey, 2006).

For both sets of four-syllable words there was a trend of improving, but unstable, accuracy for untreated words with the application of treatment. In order to examine the effects of response generalisation on untreated four-syllable words a magnitude of difference in baseline probe data to follow-up probe data was estimated using the Δ-index calculation for each set of untreated four-syllable words. For untreated 4 syl.-2nd words the calculation yielded an ES index of 1.79 and for untreated 4 syl.-3rd words yielded an ES-index of 1.32 (Pearson Education Inc, 2006) indicating a small effect size (Beeson & Robey, 2006) for both sets of words, suggesting that treatment resulted in some positive changes in sound production accuracy (generalisation) for untrained four-syllable words.

Figure 2 displays TUD (top graph) and PWC (lower graph) for treated and untreated phrases. There was only a slight decline in TUD for treated phrases with the application of treatment and post-treatment. PWC averaged 81% during baseline. Following the application of treatment, PWC increased to 90% for the final three probes of the treatment phase and tended to be on average 15% greater for treated phrases in comparison to untreated phrases. The CDC method was utilised to assist with visual inspection of this data (Figure 2, graph 2). With 10 data points in the treatment phase, the CDC method required that at least six data points in that phase be above both criterion lines to conclude there was a positive treatment effect. However, only 5 of the 10 data points in the treatment phase fell above both criterion lines, suggesting no reliable treatment effect for phrases.

Figures 3 and 4 display TUD and PWC for untrained sets of stimuli (five-syllable words, s-blend words, and sentences with four-syllable words). There was a gradual reduction in TUD over time for these stimulus generalisation items with no consistent improvement for PWC.

DISCUSSION

This study was designed to examine the effects of rate control treatment on the accuracy of sound production and total utterance duration of multisyllabic words, phrases, and sentences with an individual with AOS and aphasia. Findings revealed that treatment resulted in an improvement in sound production accuracy in an individual with AOS and aphasia. Improved sound production accuracy was observed with four-syllable words and phrases in repetition tasks.

The utilisation of CDC method indicated that positive treatment effects were obtained with four-syllable words (4 syl.-2nd and 4 syl.-3rd) and this was also supported by a large effect size for 4 syl.-2nd words, and a small effect size for 4syl.-3rd on treated words. These changes occurred even though treatment did not directly target sound production accuracy (i.e., feedback was carefully controlled so the speaker was not provided with any information regarding the accuracy of his productions). Improved sound production accuracy was observed for untrained four-syllable words, but accuracy was unstable. However, effect sizes calculated for untrained four-syllable words revealed a small effect size for both sets of words suggesting that treatment had some positive effects (i.e., generalisation) on untrained

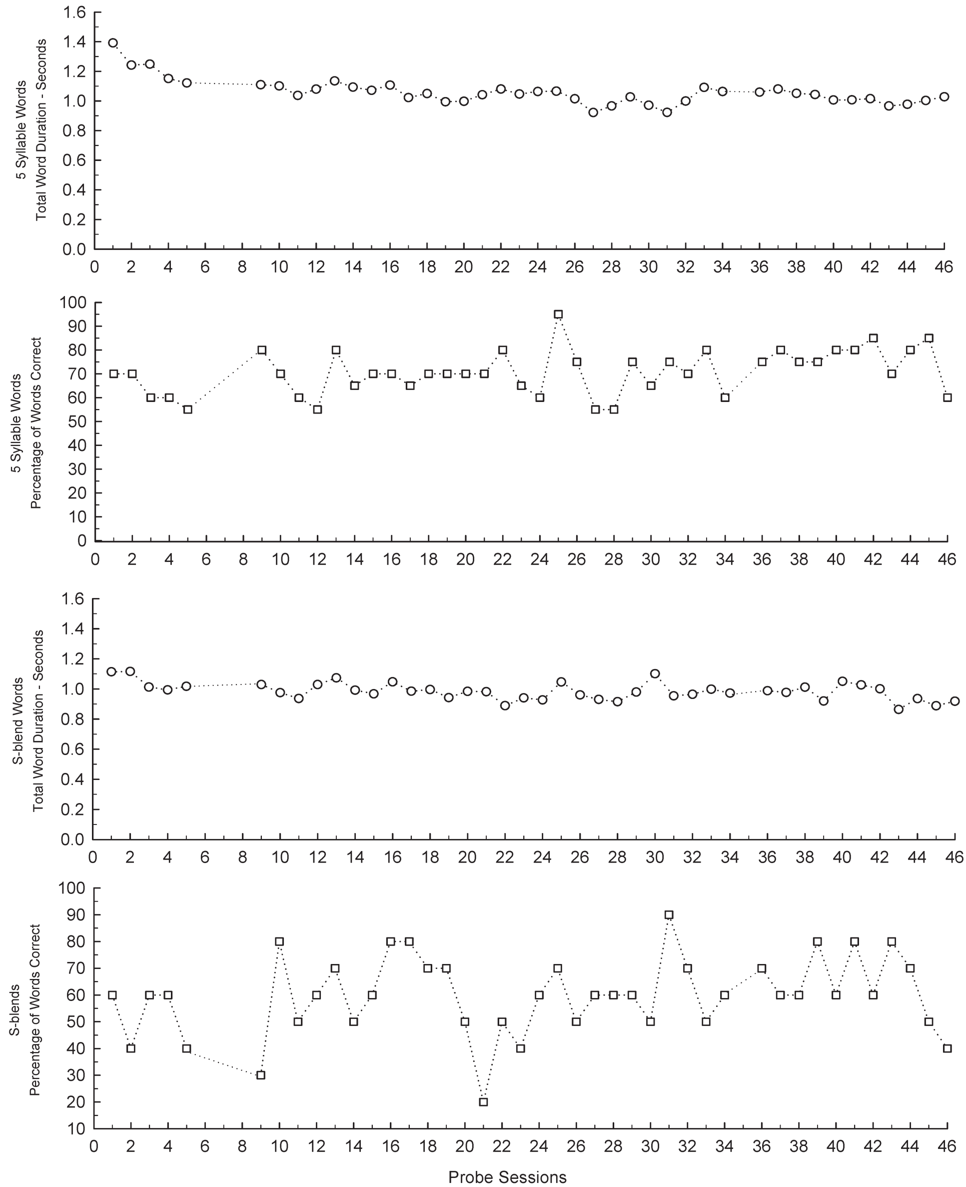


Figure 3. Total utterance duration and percentage of words correct in probes for five-syllable words (top two graphs) and s-blend words (lower two graphs).

words with increases in sound production accuracy. Phrases were selected for treatment due to the variability in sound production accuracy across baseline probes. Treatment was applied to phrases to determine if sound production accuracy would increase and stabilise for treated phrases with any generalisation to untreated phrases and/or sentence stimuli. Although there was an increase in sound production accuracy for phrases with the application of treatment, the high levels of accuracy across baseline probes (81% mean accuracy for treated phrases) likely precluded finding any treatment effect with no generalisation to untrained phrases or sentence stimuli.

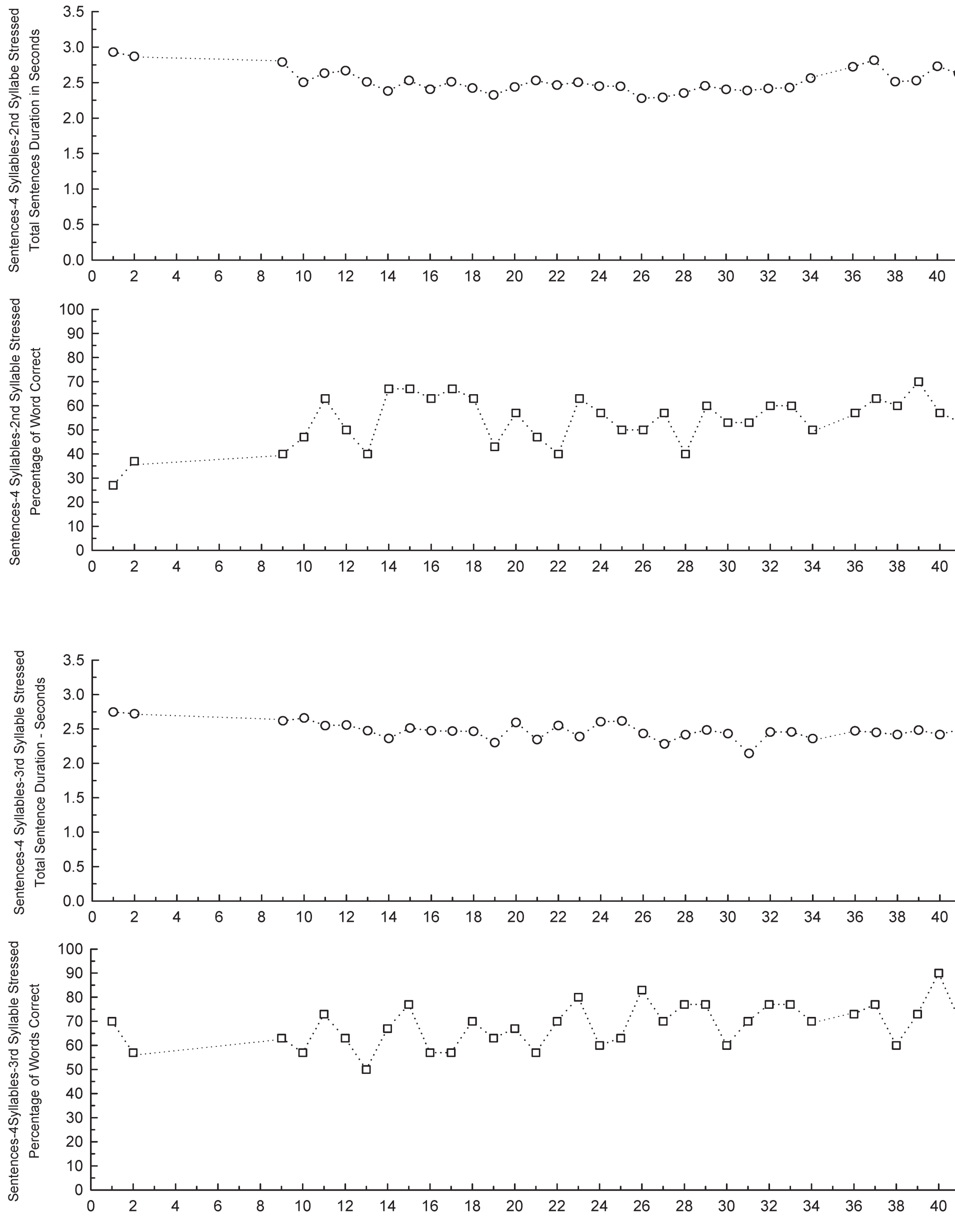


Figure 4. Total utterance duration and percentage of words correct in probes for sentences with four-syllable words with stress on the second syllable (top two graphs) and sentences with four-syllable words with stress on the third syllable (lower two graphs)

The findings in the present investigation are similar to the results of Wambaugh and Martinez (2000) with improved sound production accuracy for trained and untrained multisyllable words in a speaker with AOS and aphasia. However, Wambaugh and Martinez (2000) also had generalisation (stimulus generalisation) to untrained behaviours with an increase in articulatory accuracy for four-syllable and s-blend words, which was not found in the present investigation.

Treatment was primarily composed of repeated practice, and there is the possibility that repetition alone may have resulted in the same changes, but the positive response generalisation effects (i.e., positive changes for trained and untrained words) suggest that repetition alone was not responsible for the changes observed. Yet it could be speculated that repeated practice had a facilitatory effect on the production of syllable sequences and was the reason for the generalisation effects observed. Initially, treatment involved the clinician providing a number of models of the treatment target, which was eventually faded. It is unknown what influence clinician modelling may have had for the speaker in terms of sound production accuracy. That is, this investigation was not designed to examine components of the treatment package such as modelling and measured only the effect of the whole treatment. The impact of clinician modelling (i.e., number of models provided) in treatment should be evaluated in subsequent treatment studies.

The employment of handing tapping in conjunction with the metronome may have had a facilitatory effect on articulation with generalisation to probes. Specifically, in the initial phase of treatment there was a 10% reduction in speaking rate, which may have allowed for increased sensory feedback and provided additional motor planning time, prior to incremental increases in speaking rate. It was hoped that the utilisation of hand tapping in therapy would offer an avenue of generalisation from treatment tasks to non-treatment tasks (i.e., probe). However, the speaker was not instructed to use hand tapping outside treatment, but on occasion did attempt hand tapping during probes. Certainly, incorporating this step should be explored in future treatment studies of this kind.

There was a gradual decline in TUD over time across all stimuli with negligible differences in TUD for treated versus untreated stimuli in probes. These findings suggest that a slowed rate of production was *not* responsible for the increase in sound production accuracy. It is possible that treatment may have initially provided the opportunity for additional motor planning time and feedback, subsequently re-establishing more effective and efficient motor planning prior to incremental increases in speaking rate. However, more effective and efficient motor planning can only be inferred based on the speaker's performance of increased sound production accuracy with a higher percentage words correct for treated behaviours.

Interestingly, the speaker was easily able to increase his speaking rate during therapy tasks usually faster than targeted rate (beats per minute) while maintaining articulatory accuracy of treated words and phrases. However the speaker's ability to increase his rate of speech was not reflected in probe data. It is possible that hand tapping could also have had a facilitatory effect on the speaker's rate of production if it had been utilised outside therapy tasks. Had this step been instituted, a larger decline in utterance durations may have been seen over the course of the investigation on trained behaviours and possibly generalisation to untrained behaviours.

This study represents an initial investigation examining the effects of rate control treatment, specifically targeting an increase in rate of production on sound production accuracy and total utterance durations. Treatment resulted in improvement in sound production accuracy in word and phrase repetition tasks with an individual with AOS and aphasia, although sound production accuracy was not directly targeted during treatment. There was a decline in total utterance durations over the course of the investigation with minimal differences among treated and untreated stimuli. This type of treatment appears to have promise in terms of improving sound production accuracy and warrants further investigation. Future research should include examination of the effects of repetition versus metronome-plus repetition as well as utilising hand tapping inside and outside of therapy tasks.

Manuscript received 21 July 2007
Manuscript accepted 29 February 2008
First published online 29 April 2008

REFERENCES

Beeson, P. M., & Robey, R. R. (2006). Evaluating single-subject treatment research: Lessons learned from the aphasia literature. *Neuropsychology Review*, *4*, 161–169.

Bloom, M., Fischer, J., & Orme, J. G. (2006). *Evaluating practice – Guidelines for the accountable professional*. (5th ed.). Boston, MA: Allyn & Bacon, Pearson Education, Inc.

Brendel, B., Ziegler, W., & Deger, K. (2000). The synchronization paradigm in the treatment of apraxia of speech. *Journal of Neurolingistics*, *13*, 241–327.

Dabul, B. L. (2000). *Apraxia battery for adults* (2nd ed.). Austin, TX: Pro-Ed.

Duffy, J. R. (2005). *Motor speech disorders: Substrates, differential diagnosis, and management* (2nd ed.). St. Louis, MO: Elsevier Mosby.

Dworkin, J. P., & Abkarian, G. G. (1996). Treatment of phonation in a patient with apraxia and dysarthria secondary to severe closed head injury. *Journal of Medical Speech Language Pathology*, *4*, 105–115.

Dworkin, J. P., Abkarian, G. G., & Johns, D. F. (1988). Apraxia of speech: The effectiveness of a treatment regimen. *Journal of Speech and Hearing Disorders*, *53*, 280–294.

Fisher, W. W., Kelly, M. E., & Lomas, J. E. (2003). Visual aids and structured criteria for improving visual inspection and interpretation of single-case designs. *Journal of Applied Behaviour Analysis*, *36*, 387–406.

Kay Elemetrics. (2004). *Multi-Speech* [computer software- Model 3700]. Lincoln Park, NJ: Kay Elemetrics.

Kertesz, A. (1982). *The Western Aphasia Battery*. New York: Grune & Stratton.

McHenry, M., & Wilson, R. (1994). The challenge of unintelligible speech following traumatic brain injury. *Brain Injury*, *8*(4), 363–375.

McNeil, M. R., Robin, D. A., & Schmidt, R. A. (1997). Apraxia of speech: Definition, differentiation, and treatment. In M. R. McNeil (Ed.), *Clinical management of sensorimotor speech disorders* (pp. 311–344). New York: Thieme.

Pearson Education, Inc. (2006). *SingWin* [computer software]. Boston, MA: Allyn & Bacon, Pearson Education, Inc.

Porch, B. (2001). *Porch Index of Communicative Ability (Vol. 2): Administration, scoring and interpretation* (4^{th} ed.). Albuquerque, NM: PICA Programs.

Raven, J., Raven, J. C., & Court, J. H. (1998). *Coloured progressive matrices*. Oxford: Oxford Psychologist Press, Ltd.

Southwood, H. (1987). The use of prolonged speech in the treatment of apraxia of speech. In R. H. Brookshire (Ed.), *Clinical Aphasiology Conference proceeding* (pp. 277–278). Minneapolis, MN: BKR.

Wambaugh, J. L., & Martinez, A. L. (2000). Effects of rate and rhythm control treatment on consonant production accuracy in apraxia of speech. *Aphasiology*, *14*(8), 851–871.

Yorkston, K. M., & Beukelman, D. R. (1981). *Assessment of intelligibility of dysarthric speech*. Austin, TX: Pro-Ed.

APPENDIX: TREATMENT

Level one: Clinician model, unison production, beginning patient production

A. Schematic/tapping review.

B. Metronome setting: Initial increase in rate by 10% (bpm), then decrease in 10% increments (bpm) on two occasions.

1. Four-syllable words (primary stress on second syllable): 212, 236, 260 bpm.
2. Four-syllable words (primary stress on third syllable): 228, 248, 276 bpm.
3. Phrases (4 to 5 syllables): 179, n/a, 198 bpm.

C. Treatment steps:

1. Clinician Model (CM) – one production.
2. Patient taps along (no verbal production), while clinician produces word – three productions.
3. Unison Productions & Tapping (UPT) – three productions.
4. Patient Productions & Tapping (PPT) – five productions.

D. Target items presentation: Clinician presents experimental words (treatment items only) in random order. Clinician presents treatment items 2–4 times each session.

E. Feedback:

1. Clinician provides positive or negative feedback about tapping, production of correct number of syllables, and/or production of syllables on beat.
2. Clinician will attempt to provide no feedback regarding accuracy and inaccuracy of productions. If patient requests feedback the clinician will provide general feedback regarding accuracy or inaccuracy of production on last production during PPT.

F. Scoring:

1. "+" or "−" for first production of PPT step ("+" = correct use of tapping with production of correct number of syllables on the beat).

G. Criterion: 90% accuracy for entire treatment session in two consecutive sessions.

Level two: No clinician model, repeated patient production

A. Schematic/tapping review.
B. Metronome setting: Same as outlined above.
C. Treatment steps:

(1) Clinician says word with normal rate and prosody (no metronome or tapping).
(2) PPT – five productions.
   a) if any errors in tapping to the beat or in producing correct number of syllables – CM (one production) plus PPT (five productions); if errors remain – CM (one production) plus UPT (three productions); if errors remain – clinician presents next target word.
   b) If correct, clinician begins treatment steps with next word.

D. Target item presentation: same as level 1.
E. Feedback:

1. Clinician provides positive or negative feedback about tapping, production of correct number of syllables, and/or production of syllables on beat.
2. Clinician will attempt to provide no feedback regarding accuracy & inaccuracy of productions. If patient requests feedback the clinician will

provide general feedback regarding accuracy or inaccuracy of production on last production during PPT.

F. Scoring: "+" or "−" for first production of PPT step.
G. Criterion: 90% accuracy for entire treatment session in two consecutive sessions.

BRAIN INJURY, 1998, VOL. 12, NO. 9, 793–803

Auditory vs visual speech timing cues as external rate control to enhance verbal intelligibility in mixed spastic-ataxic dysarthric speakers: a pilot study

MARCELLA A. PILON,
KATHLEEN W. McINTOSH and
MICHAEL H. THAUT

Center for Biomedical Research in Music, Colorado State University, CO, USA

(*Received 12 January 1998; accepted 31 March 1998*)

Metronome, singing, and board pacing were used as external rate control techniques for the purpose of comparing the effectiveness of auditory and visual speech timing cues for reducing speech rate and increasing intelligibility in three traumatically brain injured mixed spastic-ataxic dysarthric speakers. A single system design with baseline reversal (ABACAD) was used in this preliminary investigation. Results demonstrated statistically significant ($p < 0.05$) changes in increased speech intelligibility during all three pacing conditions for the two more involved subjects. Differences between treatment conditions were not statistically significant. However, auditory metronome cuing showed the best results for the two subjects who benefited from rate control. Lower baseline intelligibility was strongly correlated with higher benefit from rate control. Furthermore, the two auditory rhythmic pacing conditions exhibited a close synthronization effect between the frequency rate of the cue and speech rate. Significant correlation coefficients between decreased speech rate and increased intelligibility were only found for the two more involved subjects. These findings suggested a differential benefit of slowing speech rate to improve intelligibility contingent upon severity of speech deficits.

Introduction

The presence of dysarthria as a sequelae to traumatic brain injury (TBI) is commonly reported. When it occurs, dysarthria often persists with little or no improvement up to 15 years after injury [1–4] and the effects of persisting dysarthria on speech intelligibility may be profound. Despite the potential impact of dysarthria on TBI patients, there has been relatively little research addressing the efficacy of specific intervention techniques and the appropriateness of particular techniques for certain patients [5].

Intervention strategies for improving speech intelligibility in dysarthria associated with TBI have included: the development of augmentative communication systems for unintelligible dysarthric speakers [6], restoration of the aeromechanical integrity of respiratory, laryngeal and velopharyngeal mechanisms involved in producing intelligible speech [4], treating the articulatory components of speech through normalization of abnormal oral muscular tone or compensation for motor limitations [7], and reducing speaking rate through external rate control strategies [7].

Correspondence to: Michael H. Thaut, Ph.D., Department of Music and Department of Electrical Engineering, Colorado State University, Fort Collins, CO 80523, USA.

In regard to the latter, Yorkston and Beukelman [7] and Yorkston *et al.* [8] suggest that reduction of speaking rate, using rate control techniques, is an important and one of the most effective strategies in treatment to improve intelligibility. It is thought that reducing an individual's speaking rate may improve intelligibility either because dysarthric speakers have more time available to achieve accurate articulatory targets or because the number of irregular articulatory breakdowns may be diminished during slower speaking rates. It is also suggested that reduced speaker rate may simply allow listeners more processing time thus increasing listeners' perceptions of speech intelligibility.

Yorkston and Beukelman [9] reported increases in speech intelligibility using a hierarchy of rate reduction techniques with four ataxic dysarthric speakers, two of whom were traumatically brain injured. The rate reduction strategies used by these authors provided external pacing guides for speakers to modify their rate during speaking tasks including: an alphabet board (requiring speakers to point to the first letter of each spoken word on an alphabet board as they spoke), a pacing board initially described by Helm [10] (requiring patients to point sequentially to slots on a board as words or syllables were produced), 'rhythmic cueing' (using natural speech pauses and phrasing provided by a clinician pointing to words as the speaker read them), and oscilloscopic feedback (in which speakers were required to 'fill-up' or match a 5 second window of an oscilloscopic display).

Other studies have demonstrated the usefulness of these specific techniques with dysarthric speakers. Crow and Enderby [11] reported improvements in intelligibility and articulatory accuracy in six dysarthric speakers using an alphabet board, compared to unaided oral speech. Helm [10] found that board pacing effectively reduced the speaking rate and increased the speech intelligibility of a hypokinetic dysarthric speaker who had otherwise been unable to volitionally modify speech production. The effectiveness of computer pacing has been demonstrated by Yorkston *et al.* [12] with a patient with ataxic dysarthria. This technique combines attributes of board pacing and 'rhythmic cueing', requiring speakers to read along with a cursor which underlines words in sentences as it moves across a computer screen. The rate of cursor movement and subsequent word production rate are preset at target rates according to individual patient needs.

While specific speech pacing techniques such as alphabet board, board pacing, 'rhythmic' pointing cues, and oscilloscopic feedback incorporate visuo-spatial cues as timing guides to reduce speaking rates, other techniques use auditorally presented rhythms to pace speech production. Metronome pacing is one such strategy which teaches speakers to produce words or syllables in sentences in conjunction with a regular, repetitive beat. Speaking along with the beat patterns in sung melodies may also be used as a pacing technique which provides external auditory cues to reduce speaking rates. The use of metronome and singing pacing devices has not been reported with TBI dysarthric speakers.

With evidence that rate control strategies provide an important approach to the treatment of dysarthria associated with TBI yet a lack of research comparing the effectiveness of external timing cues, particularly those which incorporate different modalities of presentation, this study compared, in a preliminary investigation, the efficacy of three speech pacing techniques, two using auditory rhythmic external timing cues [metronome pacing (MP) and singing pacing (SP)] and one using a visuo-spatial, non-rhythmic timing cue [board pacing (BP)]. Effectiveness was further assessed with regard to two issues to guide future research efforts: (1) did these techniques actually elicit designated speech rate reductions? and (2) what were the consequences of slowed speech rates using three different techniques on speech intelligibility rating for three TBI speakers with different levels of dysarthric impairment?

Method

Subjects

The subjects for this study were three male post TBI patients from an outpatient brain injury recovery programme. They were recommended for participation in this study based on speech pathology assessments of residual speech dysarthria. Each subject had sufficient visual and reading abilities, and possessed functional hearing acuity for the support of speech. In addition, each subject was at least 12 months post trauma, and had no pre-trauma neurologic or speech and language problems.

Along with the diffuse nature of TBI, the characteristics of dysarthrias following head injury often involve features of mixed dysarthria types, according to conventional dysarthria classification guidelines [7]. The subjects referred for this study also presented with mixed dysarthric features, primarily spastic-ataxic.

Subject 1 was a 23-year-old male, 24 months post TBI, with severe spastic-ataxic dysarthria characterized by hypernasality, slow laboured, imprecise and somewhat irregular articulation, monopitch and monoloudness which contributed to moderate to severe compromise in speech intelligibility. This patient also presented concomitant mild to moderate dysphagia. Neuroimaging studies reported post traumatic encephalomalacia involving temporal, frontal, and parietal regions with moderate dilation of lateral and third ventricles, and brainstem injury associated with severe TBI 2 years earlier. Double hemiparesis with spasticity (greater on the left than right), and moderate to severe diffuse cognitive impairments were also present.

Subject 2 was a 23-year-old male, 18 months post TBI, with a mild to moderate spastic-ataxic dysarthria characterized by slow imprecise articulation, monopitch, monoloudness, equal syllable stress and slight hypernasality all of which contributed to mild compromise in speech intelligibility. Early neuroimaging studies demonstrated bilateral intraparenchymal haemorrhagic lesions, mild ventriculomegaly, and brain stem injury associated with severe TBI 1.5 years earlier. This patient also presented with a moderately ataxic gait.

Subject 3 was a 44-year-old male, 12 months post TBI, with moderate to severe mixed spastic-ataxic dysarthria characterized by slow, laboured and imprecise articulation, monopitch, monoloudness, slight resonance distortion, and lingual incoordination which contributed to moderate to severe compromise in speech intelligibility. This patient also presented with left facial weakness and dysphagia. Neuroimaging studies reported deep midline cerebellar involvement and encephalomalacia of frontopontine and occipitotemporal pontine tracts extending inferiorly and along the right lateral temporal lobe. This patient also presented with severe left hemiparesis, balance problems, and bilateral upper extremity fine motor incoordination. Subject characteristics are summarized in table 1.

Table 1. Subject characteristics

Subject	Age	Dysarthria type	Level of impairment	Site of lesion	Time post onset
1	23 years	Mixed spastic-ataxic	Severe	Bilateral temporal, frontal, parietal and brainstem TBI	24 months
2	23 years	Mixed spastic-ataxic	Mild to moderate	Brainstem TBI with bilateral intraparenchymal	18 months
3	44 years	Mixed spastic-ataxic	Moderate to severe	Cerebellar, TBI, frontopontine, and occipitotemporal tracts	12 months

Design and procedure

A single system design with baseline reversal (ABACAD) was used in this study. Data were recorded on each subject's verbal intelligibility under every condition. Speech rate in words per minute (wpm) was calculated to determine if there was a correlation between speech rate and verbal intelligibility. The independent variable was the pacing technique under four conditions: no pacing (NP), singing pacing (SP), metronomic pacing (MP), and board pacing (BP). While control condition A (NP) alternated systematically with the three treatment conditions, the presentation order of the three treatment conditions was counterbalanced across subjects. Each subject participated in one session per week for a total of 6 weeks.

During NP, data were recorded for speech rate and verbal intelligibility by having the subject read a set of 30 sentence samples typed on paper. Different sentence sets were used during each baseline trial. During the treatment conditions subjects were presented with different sets of 30 sentence samples.

During SP, the investigator sang the melody at a tempo 20% slower than baseline (determined by wpm during the first baseline for each subject and expressed in beats per minute), while playing the sung melody with an accompanying harmonic chord pattern on a keyboard. The melody, tempo, stress and intonation patterns were closely related to those that may be expected form the inflection pattern and content of each sentence sample. For example, a sentence sample that was a question had a rising intonation pattern towards the end of the sentence. Next, the subject was asked to sing the sentence to the best of his ability, using the same melody, tempo, stress and intonation patterns, while the investigator accompanied the subject by singing and playing the same pattern on the keyboard. Finally, the subject was instructed to sing the sentence to the best of his ability, utilizing the same melody, tempo, stress and intonation pattern without assistance from the investigator. Data on verbal intelligibility were recorded from this last sentence. The variation in melodic patterns consisted of no more than three and a half whole-tone steps away from starting pitch and of no more than three pitch changes from initial pitch.

During MP, the subjects were instructed to read each sentence with an accompanying metronomic cue. This cue was produced via a Matrix MR-500 quartz metronome and was set at 20% below each subject's first baseline speech rate determined in wpm. The subject was instructed to verbalize each word to each beat. The subject was given an opportunity to verbalize each sentence and then go on to the next sentence sample. There was no modelling from the experimenter in this condition.

During BP, the subject was presented with one sentence at a time on a pacing board. The pacing board was approximately 13″ long and 2″ wide. Along the board, there were dividers separating spaces approximately 3″ wide where each word in the sentence sample was presented. Each divided area had a different colour which was arbitrarily chosen and was meant to make the segments more salient. Subjects were given verbal instructions and examples, followed by an opportunity to practise a sample sentence that was not part of the test before they performed the experimental sentences. Subjects were not given any rhythmic cue to facilitate slowed rate during sentence reading because the pacing board alone was considered a pacing technique. Instructions were given to tap one finger from left to right, from word to word while verbalizing each word. A similar pacing board technique has been utilized by Helm [10] and Yorkston and Beukelman [9].

Speech samples

Speech samples consisted of functional sentences chosen because they approximated the demands of ordinary speaking [12]. Thirty different functional sentences were developed for each of the six testing conditions for a total of 180 sentence samples for every subject. Sentences were selected from adult reading material and met the following criteria:

(1) Each sentence consisted of 5 to 15 words;
(2) Words in sentences occurred on a list of most frequently occurring words; and
(3) Sentences contained no numbers larger than '10', or parentheses.

The criteria used for selecting speech samples were developed by Yorkston and Beukelman [9].

To control for predictability and learning of sentences through rehearsal, no sentence was repeated within the speech sample. Each subject was presented visually with the same 180 sentence sample. Each sentence was presented one at a time on a sheet of paper directly in front of the subject. A Sony CFS-200 tape recorder and cassette tape for each subject was used to record the speech samples for all six conditions.

Data analysis

Subjects' habitual speech rates were calculated from the initial no treatment condition (baseline) using a tape recorder and a stop watch. The time from the beginning to the end of each sentence sample was determined and speech rate was calculated in words per minute (wpm): wpm = *number of words*/time (seconds) × 60. These data were used to determine the rate at which the singing and metronomic pacing techniques were implemented, as well as to compare rates of speech under each condition. A normal rate has been noted as being 190 wpm [12].

Verbal intelligibility was measured by the percentage of total words in a speech sample which were transcribed correctly. A panel of 18 judges who were unfamiliar with the sentences and who had no prior knowledge of the characteristics of dysarthric speech were chosen to listen to the recorded sentence samples. Each judge listened individually over headphones to two different subjects under two different treatment conditions, and transcribed 60 sentences word for word. This standardized technique for measuring intelligibility has been used by Yorkston and Beukelman [9] and Yorkston *et al.* [13]. The judges were instructed to listen to each sentence in its entirety before transcribing the complete sentence. The judges were not allowed to listen to any sentence more than once. Interrater reliability was computed showing a Pearson correlation coefficient of $r = 0.90$.

The percentage of total words in each sentence sample transcribed correctly was calculated for each subject. Next, the percentage of total words transcribed correctly was calculated for the entire set of 30 sentences in each testing session. Finally, the percentage of total words transcribed correctly under each testing condition was calculated and used to indicate effectiveness of each treatment and no treatment condition on verbal intelligibility.

Data were analysed visually by plotting wpm and intelligibility data in two-dimensional line graphs. Statistical analysis employed Analysis of Variance procedures (ANOVA) with planned comparisons to study the difference between treatment conditions and a Pearson Product Moment Correlation analysis to study the relationship between wpm and intelligibility scores.

Results

Mean intelligibility scores for the total study sample under the treatment and baseline conditions are summarized in figure 1 and table 2. All three pacing conditions

Table 2. Average percentage of intelligible words (%)

	No pacing	Metronome pacing (MP)	Board pacing (BP)	Singing pacing (SP)
Subject # 1	42.3 ± 13.6	62	53	59
Subject # 2	83.7 ± 10.1	83	83	75
Subject # 3	72.0 ± 1.0	94	90	88
Total average intelligibility	66 ± 20.3	80	75	74

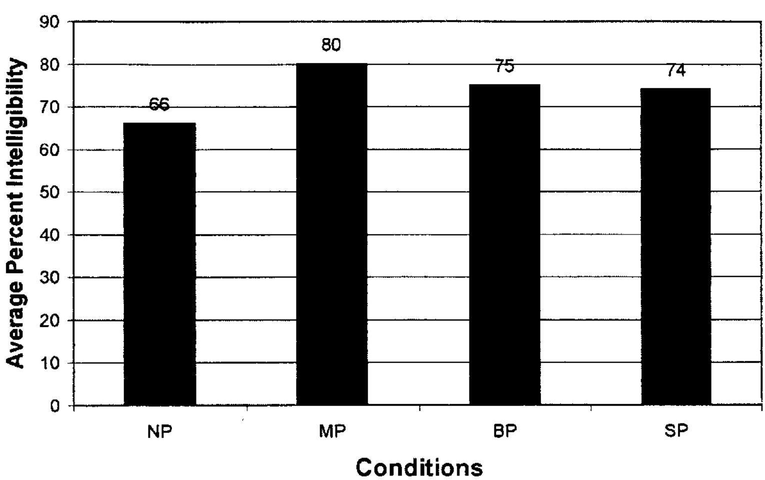


Figure 1. Mean intelligibility for three subjects across conditions.

yielded higher intelligibility scores, when averaged across subjects, than the baseline, NP (66%). Metronome pacing yielded an improvement to 80% intelligibility (21.2% increase over baseline), the pacing board resulted in an increase to 75% (13.6% over baseline), and singing showed a 74% intelligibility score (12.1% over baseline).

A review of conditions by subject scores in table 2, however, suggests differential effects of each pacing technique on improving intelligibility among the three subjects. Subjects 1 and 3 showed positive benefits from all three pacing techniques. The metronome pacing yielded the strongest improvements in intelligibility for both subjects, 46.6 and 30.6% over baseline respectively. For subject 1, singing provided the second base improvement rate (39.5% improvement, from 42.3 to 59%), and board pacing showed the smallest improvement rate (25%). The effectiveness of board pacing and singing was reversed for subject 3 who benefited slightly more from board pacing (25% improvement) than singing (22.2% improvement). Subject 2 whose baseline intelligibility score was almost twice as high as subject 1 and 14% higher than subject 3, did not benefit from any pacing condition. Board pacing and metronome each resulted in a decrease of slightly under 1%, and singing actually decreased his intelligibility by 10.4%.

Repeated measures ANOVA with planned comparisons between treatment conditions for pooled subjects, showed a difference approaching statistical difference ($F = 4.81$; $p = 0.053$) between baseline and all three pacing conditions combined. Differences between treatment conditions were not statistically significant. Within subject ANOVA procedures for each subject using the 180 trial sentences as degrees of freedom, showed a significant difference between the treatment conditions vs the no pacing baseline for subjects 1 ($F = 5.9$; $p = 0.001$) and subject 3 ($F = 9.24$; $p = 0.001$). There was no statistical difference betwen baseline and different treatment conditions for subject 2.

In addition to evidence supporting the effectiveness of external rate reduction for increasing speech intelligibility in two subjects, there was also evidence of effective rate modulation in these dysarthric speakers. Both auditory rhythmic conditions reduced the mean speech rate for all subjects by 21.6% over initial baseline. This deceleration rate was very close to the actual 20% decrease given by the rhythmic cues, indicating a synchronization effect between external cue and speech rate. Visual board pacing, without additional timing cue, also showed subjects speech rate, however at a lesser rate (11.34%) (table 3).

Figure 2 presents a visual display of percentage of intelligible words plotted against speech rate (in wpm). The graphs illustrate an inverse relationship between intelligibility and speech rate for subjects 1 and 3. For both subjects, a decrease in wpm resulted in a higher percentage of intelligible words and a return to no pacing resulted in increased wpm with decreased intelligibility. Data for subject 2 showed the oposite trend where slowed speech during each treatment condition actually

Table 3. Mean speech rate for three subjects (100% baseline)

Condition	Mean of speech rate
Metronome pacing (MP)	79.6%
Board pacing (BP)	90.3%
Singing pacing (SP)	79.6%

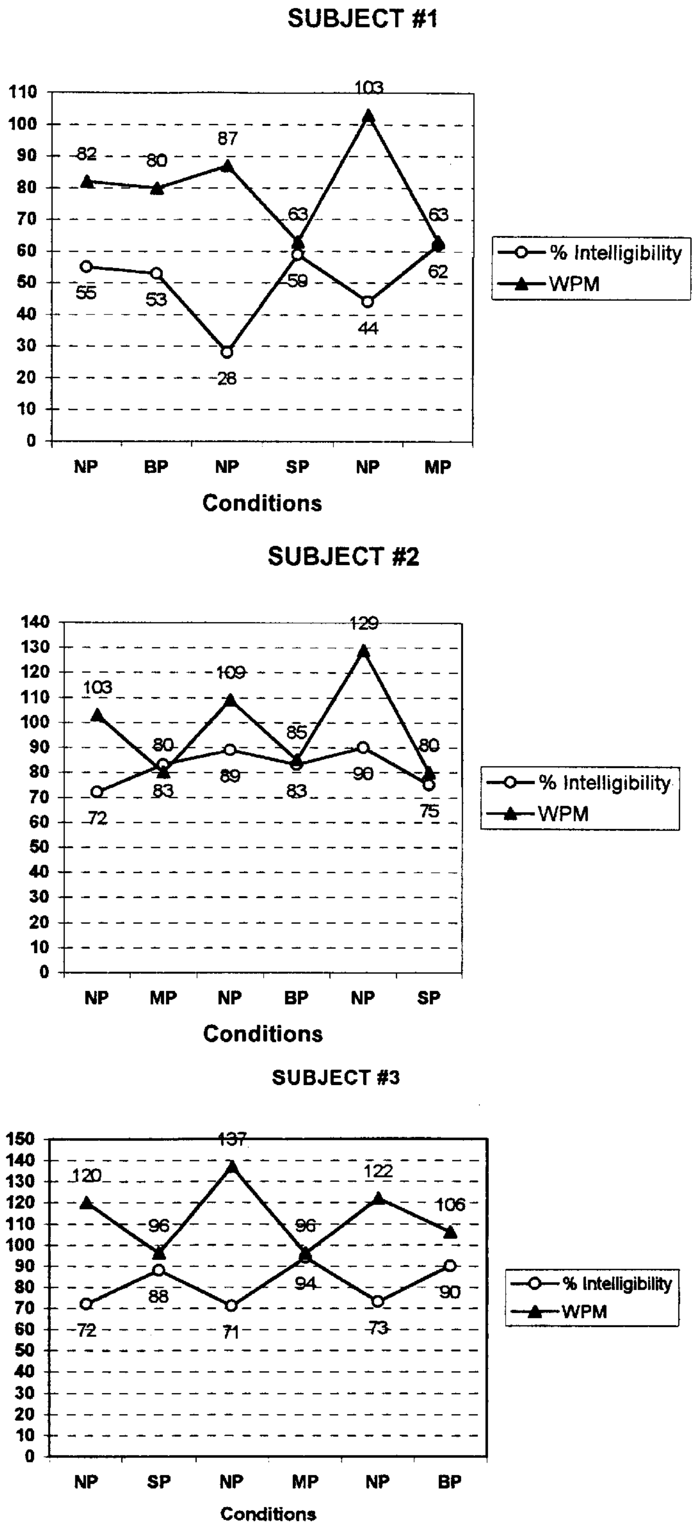


Figure 2. *Words per minute and percentage of intelligibility across contitions for individual subjects.*

was associated with a consistent yet small decrease in intelligibility and faster wpm increased intelligibility.

Pearson Production Moment correlation analysis was used to statistically assess the relationship between intelligibility and wpm. The results indicated that the inverse relationship between speech rate and intelligibility for subjects 1 and 3 was significantly correlated (subject 1: $r = -0.672$; $p < 0.05$, subject 3: $r = 0.91$; $p < 0.05$). Speech rate and increase in intelligibility were positively correlated in subject 2 without reaching statistical significance ($r = 0.52$, $p > 0.05$).

Discussion

The results of this pilot may be summarized in four major findings which may facilitate further investigation. The first three have to do with the effectiveness of the pacing techniques used in this preliminary study and the consequences of reduced speaking rate on intelligibility of three speakers with different levels of impairment.

First, the results demonstrate the effectiveness of both visuo-spatial and auditory temporal pacing cues to enhance intelligibility by decelerating speech rate in three mixed spastic-ataxic dysarthric speakers. However, individual differences in performance were most noteworthy and may offer the most useful insight into the clinical effectiveness of the treatment conditions investigated in this preliminary study. Slowing of speech rate was only successful and statistically significant in enhancing speech performance in the two subjects with the lower baseline intelligibility scores whereas the subject with an intelligibility score of over 80% did not benefit from the pacing board and metronome and actually decreased with singing. Actual mean increase in speech intelligibility due to slower pacing across all three treatment conditions showed an inverse proportional relationship with the initial baseline intelligibility score, i.e. the lower the intelligibility score was during no treatment the higher the improvement rate was during rate control treatment. The subject with the lowest intelligibility score (42.3%) increased his score by an average of 37%, the subject with an intelligibility score of 72% increased by 25.9%, and the subject with the highest speech intelligibility (83.7%) actually decreased during pacing by an average of 3.8%, mostly due to a loss of intelligibility during singing. Indeed, a *post hoc* correlational analysis showed a highly significant coefficient of 0.79 ($p = 0.001$) between initial baseline intelligibility and mean improvement rate during pacing treatment. This may suggest that more compromised speech performance may benefit from slowed speech pacing whereas at higher rates of intelligibility slowed speech pacing may no longer enhance performance. In this study this threshold of benefit occurred somewhere between 72 and 83.7% of intelligibility. Furthermore, singing as the pacing technique that was rhythmically and melodically most removed from normal speech prosody, actually seemed to interfere with the performance of the high-level subject who had a baseline intelligibility of 83.7%. A similar trend was seen in the subject with the second-highest intelligibility performance where singing was also the least effective treatment condition. In the most severely affected subject singing produced more intelligible speech than pacing board facilitation.

Second, visual and correlational analysis suggests that in the two more affected subjects improved intelligibility was indeed closely related to slowed speech rate. In the subject whose non-paced speech intelligibility was close to or above 80%, artificial slowing of speech rate from the periphery actually resulted in no further improvement. Thus, speech pace slowing may be one mechanism underlying facilitation of enhanced speech intelligibility in mixed spastic-ataxic dysarthric speakers.

Third, the consistency of the metronome to produce the best results with the two subjects who still benefited from speech pacing may suggest the particular effectiveness of rhythmic anticipatory cues to modulate speech output when timing is a variable of experimental and clinical interest. Yorkston *et al.* [12] have proposed that as the severity of speech disability increases, increasingly rigid rate control become more appropriate. With this in mind, it may be that the increasing benefit realized from the metronome for increasing degree of intelligibility deficits with moderate to severe mixed dysarthrias, was not only related to the anticipatory regularity of the metronome meter but also to the precision and rigidity of the pacing demand accomplished by the metronome, in comparison to singing and board pacing, both of which may provide more flexible demands.

Lastly, with respect to whether the pacing techniques employed in this study actually elicited designated speech rate reductions, it appears that all the three pacing techniques did elicit designated speech rates in the desired direction. Furthermore, the close rate coupling between pacing stimulus and speech rate with the metronome and singing suggests that speech rate can become entrained to an external auditory rhythm and thus rate changes may become influenced in a predictable manner by the paper frequency. The reduction of speech rate by visuo-spatial pacing board facilitation was lesser than with auditory rhythmic metronome and singing cues, and stayed closer to the subjects' habitual rate.

It is important to note that rate reduction with a pacing board is driven by internal timing by conjunction with visuo-spatial cues whereas singing and metronome utilize external auditory cuing purely in the temporal modality. Temporal rate stability therefore probably fluctuated at times during board pacing as the subject pointed with varied time intervals to each word and the spoken word did not always match the one to which they were pointing.

Conclusions

The findings of this study suggest that there may be differential benefits from speech rate modification techniques which are related to the severity of dysarthric impairment as well as to the modality and temporal regularity of external cueing used. Specifically, the performance of speakers in this investigation suggested that external pacing for the purpose of reducing speaking rate and increasing speech intelligibility may be beneficial when there is at least moderately severe impairment; but may be detrimental to overall speech intelligibility when there is only mild speech impairment. For the speakers in this investigation for whom speech rate modification was beneficial, findings suggested that auditory rhythmic cueing was preferable to visuo-spatial cues, not only for increasing speech intelligibility but also for effectively modulating speech to a target rate. Finally, this investigation offered preliminary evidence that rigid auditory rhythmic pacing cues could be more beneficial for TBI dysarthric speakers than pacing cues which provided more flexible pacing demands. While the results of this study are preliminary, it is hoped that they will set the stage for future investigation of speech rate modification in traumatically brain injured persons with persisting dysarthrias as well as provide insight for clinical practitioners regarding treatment selection.

References

1. SARNO, M., BUANAGURO, A. and LEVITA, E.: Characteristics of verbal impairment in closed head injured patients. *Archives of Physical Medicine and Rehabilitation*, **67**: 400–405, 1986.
2. THOMSEN, V.: Standardized methods of assessing and predicting outcome. In M. Rosenthal, E. Griffith, M. Bond and J. Mill (editors) *Rehabilitation of the Head Injured Adult* (Philadelphia: F. A. Davis Company), pp. 59-74, 1983.
3. YORKSTON, K. M., HONSINGER, M. J., MITSUDA, P. M. *et al.*: The relationship between speech and swallowing disorders in head injured patients. *Journal of Head Trauma Rehabilitation*, **4**: 1–16, 1989.
4. NETSELL, R. and LEFKOWITZ, D.: Speech production following traumatic brain injury: clinical and research implications. In R. Netsell Phd. (editor) *ASHA Special Interest Division 2 Newsletter*, **2**, Rockville, MD, pp. 1–8, 1992.
5. KENT, R. D.: The clinical science of motor speech disorders, a personal assessment. In J. A. Till, K. M. Yorkston and D. R. Beukelman (editors) *Motor Speech Disorders: Advances in Assessment and Treatment* (Baltimore, MD: P. H. Brookes Company), pp. 3–18, 1994.
6. DERUYTER, F. and KENNEDY, M. R. T.: Augmentative communication following traumatic brain injury. In D. R. Beukelman and K. M. Yorkston (editors) *Communication Disorders following Traumatic Brain Injury, Management of Cognitive, Language and Motor Impairments* (Austin, TX: Pro-Ed), pp. 317–365, 1991.
7. YORKSTON, K. M. and BEUKELMAN, D. R.: Motor speech disorders. In D. R. Beukelman and K. M. Yorkston (editors) *Communication Disorders following Traumatic Brain Injury, Management of Cognitive, Language and Motor Impairments* (Austin, TX: Pro-Ed), pp. 251–315, 1991.
8. YORKSTON, KJ. M., DOWDEN, P. A. and BEUKELMAN, D. R.: Intelligibility measurement as a tool in the clinical management of dysarthric speakers. In R. D. Kent (editor) *Intelligibility in Speech Disorders* (Philadelphia, PA: John Benjamins Company), pp. 245–285, 1992.
9. YORKSTON, K. M. and BEUKELMAN, D. R. Ataxic dysarthria: treatment consequences based on intelligibility and prosodic considerations. *Journal of Speech and Hearing Disorders*, **46**: 398–404, 1981.
10. HELM, N. A.: Management of palilalia with a pacing board. *Journal of Speech and Hearing Disorders*, **44**: 350–353, 1979.
11. CROW, E. and ENDERBY, P.: The effects of an alphabet chart on the speaking rate and intelligibility of speakers with dysarthria. In K. M. Yorkston and D. R. Beukelman (editors) *Recent Advances in Clinical Dysarthria* (Austin, TX: Pro-Ed), pp. 99–108, 1989.
12. YORKSTON, K. M., BEUKELMAN, D. R. and BELL, K. R.: *Clinical management of dysarthric speakers* (San Diego: College-Hill Press), 1988.
13. YORKSTON, K. M., HAMMEN, V. L., BEUKELMAN, D. R. *et al.*: The effect of rate control on the intelligibility and naturalness of dysarthric speech. *Journal of Speech and Hearing Disorders*, **55**: 550–560, 1990.

speech sounds. J Acou Soc Am 1956;28:902–905.
51. Pichney MA, Durlach NI, Braida LD. Speaking clearly for the hard of hearing: II. Acoustic characteristics of clear and conversational speech. J Speech Hearing Res 1986;29:434–446.
52. Teulings HL, Stelmach GE. Control of stroke size, peak acceleration, and stroke duration in parkinsonian handwriting. Hum Mov Sci 1991;10:315–334.
53. Yorkston K, Buekelman D, Bell K. Clinical Management of Dysarthric Speakers. Boston: Little, Brown, 1988.

Recovery from nonfluent aphasia after melodic intonation therapy:

A PET study

P. Belin; Ph. Van Eeckhout; M. Zilbovicius, MD, PhD; Ph. Remy, MD; C. François; S. Guillaume; F. Chain, MD; G. Rancurel, MD; and Y. Samson, MD

Article abstract—We examined mechanisms of recovery from aphasia in seven nonfluent aphasic patients, who were successfully treated with melodic intonation therapy (MIT) after a lengthy absence of spontaneous recovery. We measured changes in relative cerebral blood flow (CBF) with positron emission tomography (PET) during hearing and repetition of simple words, and during repetition of MIT-loaded words. Without MIT, language tasks abnormally activated right hemisphere regions, homotopic to those activated in the normal subject, and deactivated left hemisphere language zones. In contrast, repeating words with MIT reactivated Broca's area and the left prefrontal cortex, while deactivating the counterpart of Wernicke's area in the right hemisphere. The recovery process induced by MIT in these patients probably coincides with this reactivation of left prefrontal structures. In contrast, the right hemisphere regions abnormally activated during simple language tasks seem to be associated with the initial persistence of the aphasia. This study supports the idea that abnormal activation patterns in the lesioned brain are not necessarily related to the recovery process.

NEUROLOGY 1996;47:1504–1511

Functional brain imaging techniques have become an important tool in the study of normal human brain function and are now being applied to study the mechanisms of poststroke functional recovery. Such studies provide new insights into the plasticity of the mature cerebral cortex by showing that brain areas other than the ones normally used in certain cognitive tasks may be activated during functional recovery.[1-3] The current assumption is that these abnormal activations may reflect new emerging functional properties responsible for the recovery process. Nonetheless, an alternative hypothesis is that such abnormal activations could also reflect a maladaptive process and could in some cases explain the persistence of deficits. This hypothesis is difficult to test because of the intermingled recovery and residual deficit that are observed in most stroke patients during the progressive course of spontaneous recovery.

With regard to aphasia, language tasks may induce abnormal activation of right hemisphere structures in chronic aphasic patients.[4-10] Although these abnormal activation patterns are clearly consequences of the lesion, their relation to language recovery remains unclear. Thus, Knopman et al.[5] reported that the right inferior frontal region was activated when aphasic patients whose recovery was incomplete listened to words, whereas activation took place in the left posterior temporoparietal region in patients with good recovery. Conversely, prominent right hemisphere activations have also been reported in patients with good language recovery.[7,9] Recently, Weiller et al.[11] studied six patients with excellent spontaneous recovery from Wernicke's aphasia. Using pseudo-word repetition and verb generation tasks, they observed with positron emission tomography (PET) "clear right hemisphere activation in superior temporal gyrus and inferior premotor and lateral prefrontal cortices, homotopic to the left hemisphere language zones." However, important language areas were also activated in the left hemisphere of these patients. This bilateral involvement

From Service Hospitalier Frédéric Joliot (P. Belin and Drs. Zilbovicius, Remy, and Samson), DRM-CEA, Orsay; Centre du Langage (Ph. Van Eeckhout, C. François, S. Guillaume, and Dr. Chain), CHU Pitié-Salpêtrière, Paris; Inserm U316 (Dr. Zilbovicius), CHU Bretonneau, Tours; and Service des Urgences Cérébro-Vasculaires (Drs. Rancurel and Samson), CHU Pitié-Salpêtrière, Paris, France.

Supported in part by grant no. CRE-921010 from INSERM.
Received March 8, 1996. Accepted in final form April 23, 1996.

Address correspondence and reprint requests to Pascal Belin, Service Hospitalier Frédéric Joliot, CEA 4, Place du Général Leclerc, 91401 Orsay Cedex, France.

Table 1 Characteristics of the patients

	Patients						
	1	2	3	4	5	6	7
Absence of recovery (mo)	21	4	6	7	35	41	18
MIT duration (mo)	13	12	82	8	1	108	36
Verbal comprehension							
Item 02 (/72)	72/72	—/70	8/52	10/52	25/57	10/68	15/59
Item 04 (/15)	15/15	—/—	2/13	2/9	1/8	0/11	2/9
Item 05 (/12)	12/12	—/10	2/10	2/7	1/9	2/10	1/5
Verbal expression							
Item 13 (/10)	1/8	—/10	2/10	2/8	6/10	1/8	2/8
Item 14 (/14)	1/12	—/8	4/14	3/7	3/8	1/13	0/5
Item 15 (/15)	1/13	—/8	1/12	1/8	1/8	0/10	0/2
Item 19 (/105)	1/102	—/105	10/90	15/75	14/58	3/75	5/36
Aphasia severity (0–5)	1/4	—/4	0/3	0/2	1/2	0/3	0/2

Absence of recovery = delay between the stroke and the beginning of MIT (first BDAE testing); MIT duration = delay between the beginning of MIT and the PET examination (second BDAE testing); First testing/second testing = scores obtained for subtests of the BDAE on both testings. Verbal comprehension: item 02 tests verbal discrimination; item 04 tests orders execution; item 05 tests logic and reasoning. Verbal expression: item 13 tests repetition of words; item 14 tests repetition of concrete sentences; item 15 tests repetition of abstract sentences; item 19 tests denomination. Aphasia severity: low scores indicate severe aphasia. First testing was not available for patient 2, but clini-cal records indicate comparable degree of aphasia severity.

consequently makes it difficult to assess the respective contribution of right and left hemispheric activations in the patients' excellent spontaneous language recovery.

In the present study, we addressed this issue by selecting patients with a marked contrast between a poor spontaneous recovery and a subsequent dramatic language rehabilitation after melodic intonation therapy (MIT). MIT is a form of speech therapy that was structured into a formal, hierarchical treatment program in the 1970s after the observation that many aphasic patients were able to sing the words of previously learned songs better than they could speak.[12-15] Briefly, MIT consists in speaking with a simplified and exaggerated prosody, characterized by a melodic component (two notes, high and low) and rhythmic component (two durations, long and short). Some studies reported efficient MIT-induced language rehabilitation in patients with persistent aphasia, especially in nonfluent patients.[16,17] However, a recent report about MIT's current use and efficacy[18] emphasized the need for a well-designed randomized clinical validation.

We designed a PET activation study based on the opposition of normal and MIT-loaded verbal stimuli in order to distinguish residual deficit from recovery. We hypothesized that the two types of verbal tasks would result in different patterns of brain activation, and that the activations induced by normal verbal stimuli would be deficit-related, whereas those induced by MIT-loaded stimuli would be recovery-related.

Methods. *Patients.* Seven right-handed aphasic patients, aged 40 to 58 years (mean, 49.7 years), were studied. The three main inclusion criteria were (1) a unilateral left MCA infarct, (2) a persistent, severe nonfluent aphasia for months or years after the stroke (4 to 41 months; mean ± SD, 19 ± 15 months) despite repeated attempts to rehabilitate language with standard speech therapies, and (3) marked language improvement following the delayed introduction of a French version of MIT, Thérapie Mélodique et Rythmique (TMR).[19-22] Two patients had a Broca's aphasia, and five had a global aphasia. The duration of TMR therapy varied and ranged from 1 month to 9 years (37 ± 42 months).

The efficiency of language rehabilitation with TMR was assessed by comparing expression and comprehension scores on subtests of the Boston Diagnostic Aphasia Examination (French version).[23] Testing took place just before the start of TMR, and then on the day of the PET examination (table 1). Improvement with TMR was very significant in both the expression and the comprehension subtests ($p < 0.01$, paired t test).

This study was part of a project approved by the local Human Research Ethics Board, and written informed consent was obtained from all subjects.

Brain imaging protocol. Magnetic resonance (MR) and PET images were obtained on the same day for each subject. This ensured that the subject's head was positioned identically by the same investigator for both examinations. In addition, the detailed information obtained from the MR images allowed us to localize individual anatomic brain regions for the functional PET image analysis and exclude infarcted cortical areas from functional data analysis.

Anatomic images. MR images were obtained with a 0.5-tesla MR imager (MRMAX, General Electric, Milwaukee, WI). Each subject was positioned so that MR axial slices were parallel to the bicommissural line (AC-PC).[24] Skin marks were applied where indicated by the laser of the MR imager. Contiguous T_1-weighted 3-mm-thick axial slices and 5-mm-thick coronal slices were obtained throughout the brain.

Functional images. Scanning was performed using a LETI-TTV03 PET tomograph (CEA, Grenoble, France), the characteristics of which have already been described.[25] This scanner collects seven parallel 9-mm-thick transaxial slices 12 mm apart, with an in-plane resolution of 7 mm (full width at half maximum).[25] The marks drawn on the subject's face in the MR imager were used to ensure an exact repositioning of the head in the PET session. In addition, the crossed lasers permanently attached to the PET tomograph allowed the experimenter to monitor the subject's head position throughout the examination. During each PET scan, the room was darkened and quiet so that the only ambient noise was the cooling system of the tomograph. The effects of radiation attenuation by the head were corrected with the use of a transmission scan collected during exposure to a germanium 68-ring rotating source. For each condition, a regional cerebral blood flow (CBF) study (emission scan) was performed with use of a method directly derived from the $H_2{}^{15}O$ autoradiographic method, except that no arterial catheters were used.[26] An intravenous bolus of 3 ml of saline containing 2,960 to 3,700 MBq (80 to 100 mCi) of ^{15}O-labeled water was injected, and an 80-second scan was initiated when the tracer bolus entered the brain. Two-second frames were collected during the first 2 minutes following the injection. The first frame showing the arrival of radioactivity in the brain was identified. From this point, the sum of radioactive counts during the next 80 seconds was determined to generate CBF images.[27] There was a 15-minute interval between each CBF measurement.

Activation tasks. CBF was measured in four different conditions. *I. Rest:* subjects were asked to remain at rest. *II. Hearing:* subjects listened to a list of words read with a natural intonation by one of the investigators. *III. Simple repetition:* subjects heard and then repeated each word of a new list, read with a natural intonation by the same investigator. *IV. Repetition with MIT:* the investigator read the words of a new list with an MIT-like intonation, and the subjects were instructed to repeat each word with the same intonation. Words were bisyllabic and concrete, and the three lists were matched for frequency of use. Words were read at the frequency of approximately one word every 3 seconds, and if the subject did not repeat a word, the next word in the list was read 4 to 5 seconds later. In each condition, the investigator began to read the list 15 seconds before the injection. The lists of words were always read by the same investigator, i.e., the therapist who had reeducated the patients (P.V.E).

Regions of interest. A set of nine, cortical, anatomically defined regions of interest was *a priori* defined on the external surface of each hemisphere (figure 1), including the main cortical regions for language processing (Broca's and Wernicke's areas), as well as regions that have been shown to be important in hearing and repeating words, such as the mouth sensorimotor region, and Heschl's gyri. Each region was carefully defined in order to take into account the large inter-individual anatomic variability[28] and to ensure exact delimitation of the regions of interest using cerebral sulci as individual anatomic landmarks.

To define these regions, PET and MRI data were transferred to a Unix workstation (Sun Microsystems, Mountain View, CA), and the CBF images were segmented in regions of interest in three steps. (1) PET and MRI images were put in register as previously described.[29] The sum of the 3-mm-thick MRI axial slices was calculated 3 × 3 to obtain images with the same thickness as the PET slices. MRI and PET images were then superimposed with software enabling in-plane translations and rotations of isodensity contours: the optimal transformation that superimposed MRI contours on the corresponding PET images was then determined. The accuracy of this registration was assessed by checking that a successful superimposition was obtained at all brain levels. (2) A 3D anatomic reconstruction of the brain of each subject was performed with MRI data, with use of specific software (Voxtool, General Electric, Buc, France). The main cerebral sulci were then identified on the external surface of the reconstructed cortex. (3) The boundaries of the regions of interest were drawn on the MR images, according to the identified sulci, and transferred onto the PET images with use of the above transformation. The mean tissue activity was then computed for each region.

The boundaries of the lesion were clearly visible on the T_1-axial MRI slices and were used in the delineation of the regions of interest in the left hemisphere. When a region of interest was totally infarcted by MRI criteria, it was ex-

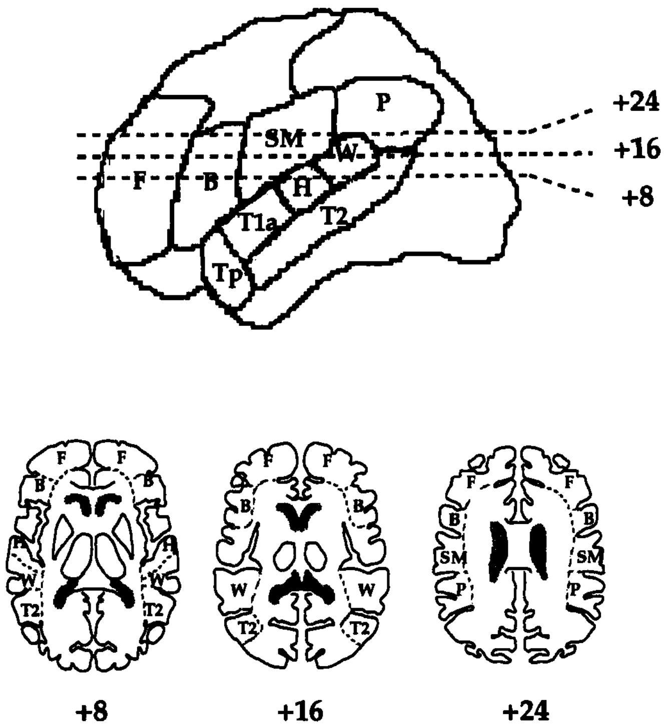


Figure 1. Regions of interest. Simplified diagrams of the nine cortical, anatomically defined, regions of interest per hemisphere. Upper: schematic lateral view of the left hemisphere. Lower: three axial slices corresponding to +8, +16, and +24 mm above AC-PC line in Talairach's space. B = Broca's area (and right hemisphere homologue). F = prefrontal area. Tp = temporal pole. T1a = anterior superior temporal gyrus. T2 = middle temporal gyrus. H = Heschl's gyri. W = Wernicke's area (and right hemisphere homologue). P = parietal area. SM = mouth sensorimotor area. See text for complete definition of the regions of interest.

cluded from the data analysis. When a region of interest was only partially touched by the lesion, the infarcted part was not further considered, and only the intact part of the region was delineated.

The anatomic boundaries of the regions of interest are detailed below.

Broca's area (B) (and its right hemisphere homologue): pars opercularis and pars triangularis of the third frontal gyrus. The upper and lower limits were the inferior frontal sulcus, and the lateral sulcus. The posterior and anterior limits were the inferior precentral sulcus and the horizontal branch of the lateral sulcus. This region was delineated on three PET slices.

Wernicke's area (W) (and its right hemisphere homologue): posterior section of the superior temporal gyrus (STG). The upper and lower limits were the end of the lateral sulcus and the superior temporal sulcus. The anterior and posterior limits were the most posterior transverse temporal sulcus and the superior temporal sulcus. This region was delineated on two PET slices.

Heschl's gyri (H): transverse temporal gyri, part of the upper bank of the superior temporal gyrus. The upper and lower limits were the lateral sulcus and the superior temporal sulcus. The anterior and posterior limits were the transverse temporal sulci. This region was delineated on one PET slice.

Anterior superior temporal gyrus (T1ant): part of the superior temporal gyrus anterior to the gyri of Heschl and posterior to the temporal pole. The upper and lower limits were the lateral sulcus and the superior temporal sulcus. The posterior limit was the most anterior transverse temporal sulcus. This region was delineated on two PET slices.

Middle temporal gyrus (T2). The upper and lower limits were the superior temporal and the inferior temporal sulci. This region was delineated on three PET slices and extended longitudinally from the end of the temporal pole to the end of the superior temporal sulcus.

Temporal pole (TP). Anterior end of the temporal lobe, where the three temporal gyri join together. The upper limit was the lateral sulcus. This region was delineated on two PET slices.

Sensorimotor mouth region (SM). This region included both pre- and postcentral gyri, and the anterior and posterior limits were the precentral and postcentral sulci. The lower and posterior limits were determined according to the level of the slice above the AC-PC line. This region was delineated on two to three PET slices, i.e., +20 and +45 mm above the AC-PC line.

Parietal region (P). This region included both supramarginal and angular gyri. The upper limit was the interparietal sulcus, and the lower limit was at the level of the upper end of the lateral sulcus. The anterior and posterior limits were the postcentral and the interparietal sulci. This region was delineated on three PET slices.

Prefrontal region (F). This large region included both first and second frontal gyri, and its posterior limit was formed by Broca's area. It extended vertically from 0 to 40 mm above the AC-PC line, corresponding to three to four PET slices.

Data analysis. Because the change in local tissue radioactivity is linearly related to blood flow under the conditions of this study, relative changes in tissue activity were taken to indicate relative changes in blood flow.[26] To correct for intercondition changes in the global activity of the brain, all PET images were normalized so that for each emission scan, each pixel of each slice was divided by the mean pixel value of the seven slices. For each subject and for each included region of interest, the relative CBF variations between conditions were computed ($\Delta\% = 100 \times$ [condition b − condition a]/condition a). These relative CBF variations were analyzed in a hierarchical manner.[30,31] First, the listening condition was compared to the rest condition, then the simple repetition condition to the listening condition, and finally a comparison was made of both repetition conditions with and without MIT. These relative CBF variations were tested for significance across the population using nonparametric Wilcoxon's signed-rank test.

Hypoperfusion in the left hemisphere was determined using rest relative CBF asymmetries. An asymmetry index ($200 \times$ [right − left]/[right + left]) exceeding 20% for a given region of interest was considered as being indicative of a significant hypoperfusion of that region in the left hemisphere.

Results. *Anatomic cortical damage.* For the purpose of the study, we carefully analyzed the extent of the lesion on MRI. All patients had an infarct in the left MCA territory, extending in two cases (patients 1 and 3) to the ACA territory. Figure 2 shows that these infarcts were large and always associated with ipsilateral dilatation of the ventricular system and thalamic atrophy. Yet, the cortical damage was comparatively less extensive than the subcortical damage. Hence, only three regions in the left hemisphere had to be excluded from the data analysis, because they were spared in less than five out of seven patients: Heschl's gyrus, which was totally infarcted in five patients, and the temporal pole and the mouth sensorimotor cortex, which were infarcted in four patients. Interestingly, Broca's area was infarcted in only two patients, and Wernicke's area in only one patient. Thus, by MRI criteria, essential language cortical areas were spared in most cases on MRI and were included in the data analysis.

Functional cortical damage. The asymmetry in rest CBF values was used to determine the extent of the hypoperfused area in the left hemisphere. For all but one patient, the hypoperfused area in the left hemisphere was very broad, extending to the whole perisylvian area, and to the prefrontal cortex in five patients. In these patients, the functional abnormality in the left hemisphere greatly exceeded the extent of structural damage in the cortex, as is generally the case in aphasic patients.[32] For the remaining patient (no. 1), the only hypoperfused region was Broca's area, which was also functionally compromised in all the other patients.

Activation study. Subject performance during both repetition tasks was variable. However, significantly more words ($p < 0.03$, Wilcoxon's rank sum test) were correctly repeated with MIT (16.3 ± 8 words) than without MIT (12.4 ± 8 words).

The comparison between tasks highlighted the following relative CBF changes (table 2, figure 3).

I. Hearing Words vs. Rest. This listening task significantly ($p < 0.02$) increased relative CBF in the entire right superior temporal gyrus: right anterior STG, right Heschl's gyri, as well as the right homologue of Wernicke's area. Relative CBF was also significantly increased in the

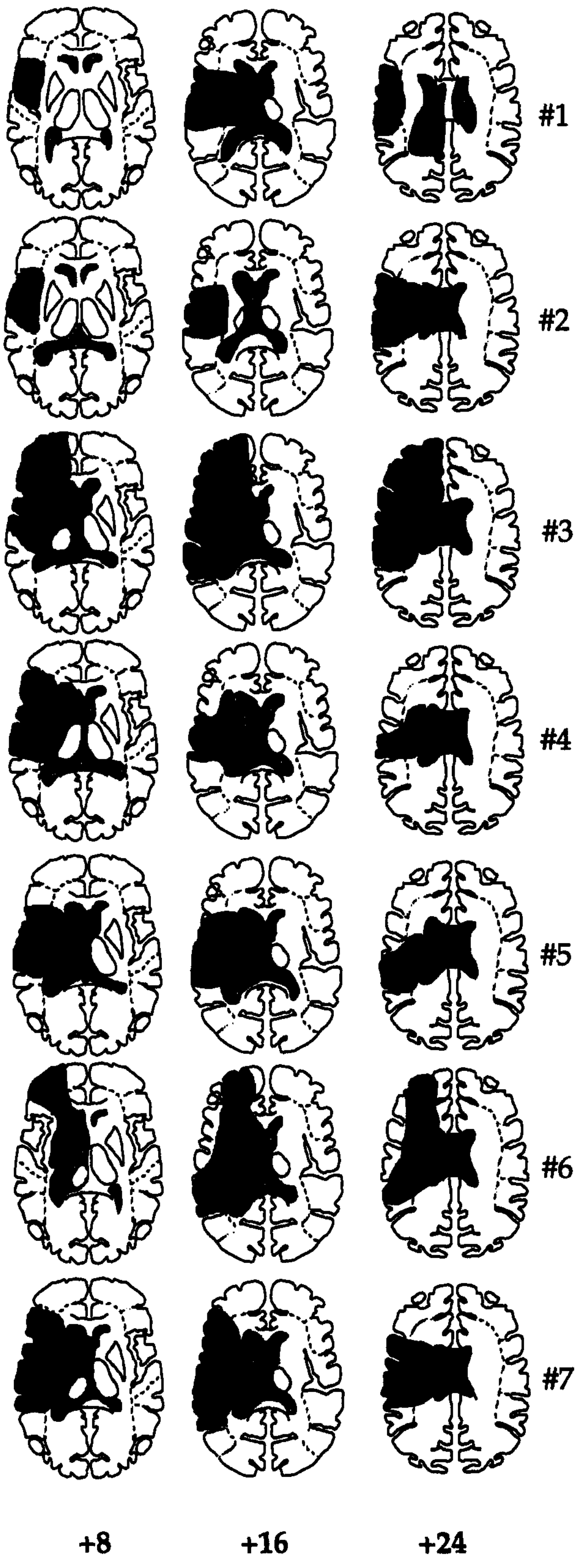


Figure 2. Extent of MRI lesions. Three axial slices corresponding to +8, +16, and +24 mm above AC-PC line in Talairach's space are given for each patient. Infarcted tissues are represented in black, ventricles in gray. The extent of the anatomic damage is variable; however, Broca's and Wernicke's areas are in most cases spared on MRI.

right prefrontal region. In the left hemisphere, only Wernicke's area showed relative CBF increase ($p < 0.04$) and this was less than its homologue in the right hemisphere.

II. Repetition without MIT vs. Hearing Words. This comparison showed significant relative CBF increases ($p < 0.02$) in the right sensorimotor mouth region and in the right hemisphere homologue of Wernicke's area. Relative CBF in the right prefrontal and anterior STG regions, which was decreased by the hearing task, was significantly increased ($p < 0.02$). In the left hemisphere, the only significant change across subjects was the decrease in relative CBF in Broca's area ($p < 0.04$).

III. Repetition with MIT vs. repetition without MIT. Introducing MIT in the repetition condition resulted in a relative CBF decrease in seven out of nine right hemisphere regions of interest, and for the group of subjects this was significant for the right homologue of Wernicke's area ($p < 0.02$). In the left hemisphere, there was a significant relative CBF increase in Broca's area, and in the adjacent prefrontal cortex ($p < 0.04$) (see figure 3).

Discussion. Two main findings emerged from this study. First, simple passive (word hearing) and active (word repetition) verbal tasks performed without MIT resulted in abnormal activation of right hemisphere structures, homotopic to those normally activated in the intact left hemisphere. Second, word repetition performed with MIT reactivated Broca's area and the adjacent left prefrontal cortex. This study has obvious limitations. For example, as shown in table 1, the duration of MIT varied widely. Yet, in view of the clear dissociation between the lack of spontaneous recovery and the subsequent dramatic effect of MIT, these results suggest that abnormal activation patterns are not necessarily related to functional recovery, but may be involved in the persistence of aphasia. They also suggest that the use of MIT coincided with a relative normalization of the activation pattern with reactivation of essential language areas.

Abnormal activation patterns during verbal tasks in aphasia: Maladaptive or well-adapted functional reorganization? The different patterns of activation found with and without MIT when our patients performed verbal tasks suggest that the right and left hemispheres were not equally involved in the recovery. When they performed simple verbal tasks *without* MIT, there was abnormal activation of right hemisphere structures, along with deactivation of left hemisphere structures. As in the study of Weiller et al.,[11] the abnormal right hemisphere activation occurred in regions homotopic to the left hemisphere language zones normally involved in those tasks. For example, during word hearing, activation occurred in both superior temporal gyri as in control subjects, but more extensively and more significantly in the right than in the left hemisphere, whereas the reverse pattern usually occurs in normal subjects.[30,33-35] During word repetition, the activated areas were centered on the mouth region of the right sensorimotor cortex, whereas activation is normally bilateral in intact control subjects.

***Table 2** Relative CBF changes in the regions of interest*

Region of interest	Hearing vs. rest		Usual repetition vs. hearing		Repetition with MIT vs. usual repetition	
	Mean	±SD	Mean	±SD	Mean	±SD
Broca's area (B)						
L	−3.5	5	**−3.9***	**2**	**4.7***	**1**
R	0.8	3	−1.9	5	−0.4	3
Prefrontal (F)						
L	−1.1	5	−2.2	10	**6.3***	**5**
R	**2.5†**	**2**	**−4.5†**	**5**	1.0	6
SM mouth (SM)						
L‡	−3.2	2	2.5	6	4.8	11
R	0.6	3	**7.1†**	**4**	−2.6	4
Parietal (P)						
L	4.1	6	−3.7	5	2.9	6
R	−1.3	3	−0.5	5	−0.6	5
Wernicke's area (W)						
L	**4.7***	**5**	−1.9	6	4.6	6
R	**7.3†**	**4**	**2.2†**	**2**	**−3.8†**	**3**
Heschl's gyrus (H)						
L‡	7.9	1	−9.2	4	5.7	3
R	**7.2†**	**3**	−0.8	6	−1.1	4
Ant. STG (T1a)						
L	5	7	1.1	7	0.6	6
R	**10.5†**	**7**	**−5.5†**	**5**	0.8	6
MTG (T2)						
L	−0.9	6	−1.2	7	4.7	15
R	−1.2	4	−1.9	6	−1.0	6
Temporal pole (Tp)						
L‡	2.8	9	−5.0	2	−1.3	15
R	2.7	5	−3.9	12	−1.3	15

Relative CBF changes obtained when comparing the four activation conditions in a stepwise way. Mean and standard deviation (SD) are in % of reference relative CBF. Significant changes are in bold.

* $p < 0.04$ (Wilcoxon's rank sum test).
† $p < 0.02$ (Wilcoxon's rank sum test).
‡ Regions of the left hemisphere that were not statistically tested, because lesioned in more than two patients.

That the right hemisphere activation observed with usual language tasks was involved in the recovery process is doubtful, for two main reasons. First, the patients remained severely aphasic during the pre-MIT period, although they had been repeatedly exposed to such usual verbal tasks during unsuccessful non-MIT speech therapy sessions and daily life. Second, no region of the left hemisphere was activated during the simple repetition task, which is in sharp contrast with the activation of both Broca's area and the left prefrontal cortex during the repetition task in the patients of Weiller et al.[11] In our group, Broca's area was conversely deactivated. This reinforced the functional CBF asymmetry detected at rest, an abnormality that is strongly linked to the severity of the aphasia.[32,36] Therefore, the abnormal deactivation of Broca's area may have contributed to the persistence of aphasia.

Moreover, during repetition with MIT, i.e., the very process associated with language recovery in these patients, Broca's area was activated as well as the adjacent prefrontal cortex in the left hemisphere, whereas the counterpart of Wernicke's area in the right hemisphere was deactivated. In consequence, as discussed below, the main effect of MIT was to correct the abnormally right-shifted pattern of activations, and to reactivate functionally impaired essential motor language zones.

Functional mechanism of MIT. MIT is assumed to work by developing right hemisphere language

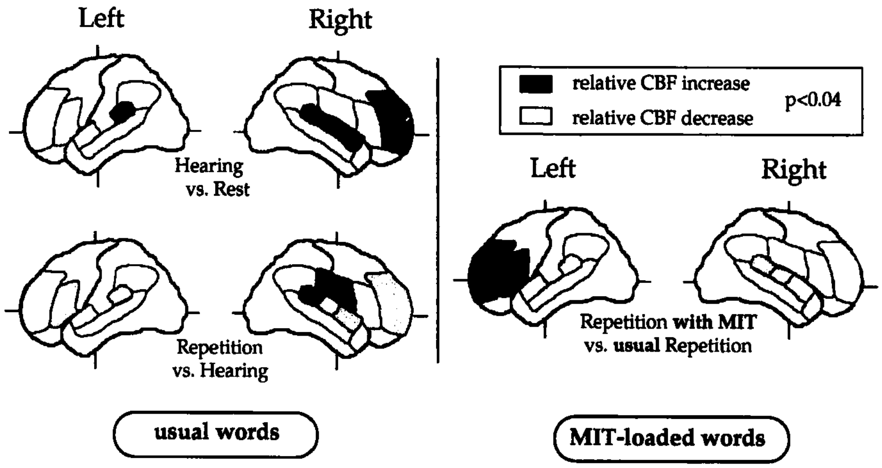


Figure 3. Significant relative CBF changes. For each comparison, regions of interest with significant (p < 0.04) relative CBF changes are represented on the left and right external views of the brain. Relative CBF increases are in black, decreases are in gray.

abilities, since the method is based on the observations of the preservation of singing abilities in severe nonfluent aphasic patients, which are at least partially mediated by right hemisphere structures.[4,31,37,38] Thus, the present results may at first seem unexpected and counterintuitive, since, as described earlier, MIT reactivated essential motor language zones, such as Broca's area in the left hemisphere, while reducing abnormal activations in the right hemisphere. A possible explanation of this apparent paradox is that MIT is not singing but merely exaggerating speech prosody. Although we do not currently understand the underlying reasons, several developmental observations suggest that exaggerating speech prosody may facilitate the recruitment of language-related brain areas. For example, mothers spontaneously exaggerate the prosodic contents of their language when speaking to young childrens.[39] Similarly, children exaggerate speech prosody when they learn new verbal material, such as poetry or multiplication tables. Exaggerating speech prosody is a complex process that alters many acoustic features of language. It is obtained in MIT through codified control of the pitch and duration of each syllable, and of the phrase rhythm.[14] Pitch control is probably not involved in the activation of Broca's area, since pitch processing activates a right hemisphere cortical network in the normal brain.[31] Syllable duration may be a more relevant parameter. Tallal's group[40-43] has shown that dysphasic children and aphasic adults are selectively impaired in the processing of rapidly changing formant transitions, and that performance is significantly improved by increasing the transition duration. Furthermore, one clinical study showed that increased syllable duration is important for MIT efficiency.[15] Finally, phrase rhythm control may also be an important parameter, since recent studies found selective activation of Broca's area by rhythm processing in normal subjects.[44,45]

Conclusions. Several findings emerged from this study. First, in line with previous results, we confirmed that language tasks may induce highly abnormal patterns of activation/deactivation in aphasic patients. Second, we observed that the behavioral correlates of these abnormal patterns are not straightforward; the abnormal patterns in some cases coincide with the persistence rather than the recovery of the aphasia. The activation of very similar right hemisphere regions seems to have contributed to recovery in the patients of Weiller et al.,[11] whereas the activation pattern in our patients seems to be maladaptive and to contribute to the persistence of aphasia. This raises the possibility that abnormal activation patterns may be a direct consequence of the lesion itself, rather than secondary phenomena indicating cortical reorganization. In this case, behavioral correlates could depend on many lesion or nonlesion factors, such as the degree of the lateralization of language brain systems before the stroke, or the size of the lesion. Third, and perhaps most important, abnormal activation patterns may possess remarkable plasticity, since they can be modified by a speech therapy method based on systematic changes in the acoustic features of perceived and produced speech.

Acknowledgments

The authors thank M.-C. Masure and A. Verwilghen for their valuable help, and M. Crouzel, H. De la Caffinière, B. Martins, and the Orsay chemical staff for technical assistance. The authors are deeply indebted to K.L. Moya and K. Parvis for critical reading of the manuscript.

References

1. Chollet F, DiPiero V, Wise RJS, Brooks DJ, Dolan RJ, Frackowiak RSJ. The functional anatomy of motor recovery after stroke in humans: a study with positron emission tomography. Ann Neurol 1991;29:63–71.
2. Weiller C, Chollet F, Friston KJ, Wise RJS, Frackowiak RSJ. Functional reorganization of the brain in recovery from striatocapsular infarction in man. Ann Neurol 1992;31:463–472.
3. Weiller C, Ramsay SC, Wise RJS, Friston KJ, Frackowiak RSJ. Individual patterns of functional reorganization in the human cerebral cortex after capsular infarction. Ann Neurol 1993;33:181–189.
4. Jeannerod M, Hecaen H. Restauration des fonctions du langage. Villeurbanne: Simep, 1979.
5. Knopman DS, Rubens AB, Selnes OA, Klassen AC, Meyer MW. Mechanisms of recovery from aphasia: evidence from serial Xenon 133 cerebral blood flow studies. Ann Neurol 1984;15:530–535.

6. Cappa SF, Vallar G. The role of the left and right hemispheres in recovery from aphasia. Aphasiology 1992;6:359–372.
7. Rousseaux M, Steinling M, Mazingue M. Activation du débit sanguin cérébral droit par le langage après une aphasie globale. Rev Neurol 1992;148:229–231.
8. Gainotti G. The riddle of the right hemisphere's contribution to the recovery of language. Eur J Disord Commun 1993;28: 227–246.
9. Price C, Wise R, Howard D, Warburton E, Frackowiak RSJ. The role of the right hemisphere in the recovery of language after stroke [abstract]. J Cereb Blood Flow Metab 1993; 13(suppl 1):S520.
10. Silvestrini M, Troisi E, Matteis M, Cupini LM, Caltagirone C. Involvement of the healthy hemisphere in recovery from aphasia and motor deficit in patients with cortical ischemic infarction: a transcranial doppler study. Neurology 1995;45:1815–1820.
11. Weiller C, Isensee C, Rijntjes M, et al. Recovery from Wernicke's aphasia: a positron emission tomographic study. Ann Neurol 1995;37:723–732.
12. Albert ML, Sparks RW, Helm NA. Melodic intonation therapy for aphasia. Arch Neurol 1973;29:130–131.
13. Sparks R, Helm N, Albert M. Aphasia rehabilitation resulting from melodic intonation therapy. Cortex 1974;10:303–316.
14. Sparks RW, Holland AL. Method: melodic intonation therapy for aphasia. J Speech Hear Disord 1976;41:287–297.
15. Laughlin SA, Naeser MA. Effects of three syllable durations using the melodic intonation therapy technique. J Speech Hear Res 1979;22:311–320.
16. Popovici M, Mihailescu L. Melodic intonation in the rehabilitation of Romanian aphasics with bucco-lingual apraxia. Rev Roum Neurol Psychiat 1992;30:99–113.
17. Naeser MA, Helm-Estabrooks N. CT scan localization and response to melodic intonation therapy with nonfluent aphasia cases. Cortex 1985;21:203–223.
18. Benson DF, Dobkin BH, Rothi LJG, Helm-Estabrooks N, Kertesz A. Assessment: melodic intonation therapy. Neurology 1994;44:566–568.
19. Van Eeckhout P, Pillon B, Signoret J-L, Deloche G, Seron X. Rééducation des réductions sévères de l'expression orale: la "thérapie mélodique et rythmée." In: Mardaga P, ed. Rééduquer le cerveu. Bruxelles: Solédi, 1982:109–121.
20. Van Eeckhout P, Honrado C, Bhatt P, Deblais J-C. De la thérapie Mélodique et Rythmée et de sa pratique. Rééduc Orthophon 1983;21:305–316.
21. Van Eeckhout P, Bhatt P. Rythme, intonation, accentuation: le rééducation des aphasies non-fluentes sévères. Rééduc Orthophon 1984;22:311–327.
22. Neumeister C, Wioland F, Metz-Lutz N, Dziony M-C, North P. Adaptation de la Melodic Intonation Therapy au rythme du français parlé. Rééduc Orthophon 1983:21.
23. Mazaux JM, Orgogozo JM. HDAE-F. Boston Diagnostic Aphasia Examination. Issy les Moulineaux: Editions Scientifiques et Psychologiques, 1981.
24. Talairach J, Tournoux P. Co-planar stereotaxic atlas of the human brain. New York: Thieme, 1988.
25. Trebossen R, Mazoyer B. Count rate performances of TTV03: the CEA-LETI time-of-flight positron emission tomograph. IEEE Trans Med Ima 1991;10:261–266.
26. Fox PT, Mintun MA, Raichle ME, Herscovitch P. A noninvasive approach to quantitative functional brain mapping with $H_2{}^{15}O$ and positron emission tomography. J Cereb Blood Flow Metab 1984;4:329–333.
27. Petit L, Orssaud C, Tzourio N, Salamon G, Mazoyer B, Berthoz A. PET study of voluntary saccadic eye movements in humans: basal ganglia-thalamocortical system and cingulate cortex involvement. J Neurophysiol 1993;69:1009–1017.
28. Steinmetz H, Seitz RJ. Functional anatomy of language processing: neuroimaging and the problem of individual variability. Neuropsychologia 1991;29:1149–1161.
29. Remy P, Zilbovicius M, Leroy-Willig A, Syrota A, Samson Y. Movement- and task-related activations of motor cortical areas: a positron emission tomographic study. Ann Neurol 1994; 36:19–26.
30. Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. Positron emission tomographic studies of the cortical anatomy of single word processing. Nature 1988;331:585–589.
31. Zatorre RJ, Evans AC, Meyer E, Gjedde A. Lateralisation of phonetic and pitch discrimination in speech processing. Science 1992;256:846–849.
32. Metter EJ. Behavioral correlates of abnormalities of metabolism in stroke patients. In: Bornstein RA, Brown G, eds. Neurobehavioral aspects of cerebrovascular disease. New York: Oxford University Press, 1991.
33. Larsen B, Skinhøj E, Lassen NA. Variations in regional cortical blood flow in the right and left hemispheres during automatic speech. Brain 1978;101:193–209.
34. Wise R, Hadar U, Howard D, Patterson K. Language activation studies with positron emission tomography: exploring brain functional anatomy with positron tomography. Chichester: Wiley, Ciba Foundation Symposium, 1991:218–234.
35. Nishizawa Y, Olsen TS, Larsen B, Lassen NA. Left-right cortical asymmetries of regional cerebral blood flow during listening to words. J Neurophysiol 1982;48:458–466.
36. Heis W-D, Pawlik G, Dietz E. Hypometabolism and functional recruitment as related to prognosis in poststroke aphasia [abstract]. J Cereb Blood Flow Metab 1991;11(suppl 2):S660.
37. Zatorre RJ, Evans AC, Meyer E. Neural mechanisms underlying melodic perception and memory for pitch. J Neurosci 1994; 14:1908–1919.
38. Blumstein S, Cooper WE. Hemispheric processing of intonation contour. Cortex 1974;10:146–158.
39. Lieberman P. Uniquely human. Cambridge, MA: Harvard University Press, 1991.
40. Schwartz J, Tallal P. Rate of acoustic change may underlie hemispheric specialization for speech perception. Science 1980;207:1380–1381.
41. Tallal P, Piercy M. Developmental aphasia: rate of auditory processing and selective impairment of consonant perception. Neuropsychologia 1974;12:83–93.
42. Tallal P. In the perception of speech time is of the essence. In: Buzsaki G, Llinas RR, Singer W, Berthoz A, Christen Y, eds. Temporal coding in the brain. Berlin: Springer-Verlag, 1994: 291–299.
43. Tallal P, Miller S, Bedi G, et al. Language comprehension in language learning impaired children improved with acoustically modified speech. Science 1996;271:81–84.
44. Platel H. Les composantes structurales de la perception musicale: étude en tomographie d'émission de positons. Caen, France: Caen University; 1995. PhD thesis.
45. Fiez JA, Raichle ME, Tallal PA, Petersen SE. Activation of a left frontal area near Broca's area during auditory detection and phonological access task [abstract]. J Cereb Blood Flow Metab 1993;13(suppl 1):S504.

APHASIOLOGY, 2003, *17* (1), 75–95

Melodic intonation therapy in Persian aphasic patients

Borna Bonakdarpour, Anita Eftekharzadeh, and Hassan Ashayeri
Teheran University of Medical Sciences, Tehran, Iran

Background: Melodic intonation therapy (MIT) is a well-known method of aphasia rehabilitation using prosodically based melodic phrases. The literature includes MIT adapted to many languages and efficiently applied to certain groups of non-fluent aphasic patients. However, there has been no report on the efficacy of the method for Persian-speaking patients.
Aims: The objective of this study was primarily to investigate the effects of 15 sessions of MIT treatment (adapted to the Persian language) in selected patients with non-fluent aphasia; primarily on expository speech (phrase length and number of correct content units) and oral expression skills (repetition, responsive naming, and confrontational naming), and secondarily, on auditory comprehension abilities (word discrimination, commands).
Methods and Procedures: Participants in the study included seven right-handed Persian-speaking patients afflicted with chronic (>14 months post-onset) non-fluent aphasia. Based on the rules of Persian prosody, MIT was adapted to the Persian language. Using a pre, post treatment design, each outcome measure was tested twice before and twice after MIT treatment. Changes in the variables not treated were also measured as baselines controlling treatment effects.
Outcomes and Results: Using the Wilcoxon signed-rank test, improvements in the selected variables were shown to be statistically significant after the treatment phase and not during the treatment-free phases. Non-target variables remained unchanged after the treatment and throughout the non-treatment phases.
Conclusions: Our study showed that MIT can be adapted for Persian aphasic patients and administered with measurable positive results after 15 sessions of treatment. MIT improved primarily spontaneous speech production, and as minor effects, selected oral expression and auditory comprehension subtests. Therefore MIT might be considered as a method for the rehabilitation of selected non-fluent Persian aphasic patients. More long-term follow-up studies with randomised controlled clinical trials are needed for stronger conclusions.

Address correspondence to: Borna Bonakdarpour, Aphasia and Neurolinguistics Laboratory, Communicative Sciences and Disorders, Northwestern University, Evanston, IL 60208-3540, USA.
Email: borna@northwestern.edu

This article was based on the authors' doctoral dissertation submitted to the Tehran University School of Medicine, February 2000.

We would like to thank Dr Jamsheed Lotfi, Ms Ariaii, Ms Abbasi, and Mr Khamooshi for their general support of the study, Dr Hassan Riahi who kindly aided us in finding ways to arrange recitative style for Persian verses, and Dr Shahram Yazdani for his helpful suggestions regarding statistical analysis of the results. We are also extremely grateful to Dr Cynthia K. Thompson who kindly reviewed the final draft of this paper.

The authors are deeply indebted to Professor Nancy Helm-Estabrooks without whose valuable guidance this study would not have been possible.

http://www.tandf.co.uk/journals/pp/02687038.html DOI:10.1080/02687030244000464

INTRODUCTION

Melodic intonation therapy (MIT) is now a well-known treatment method for the rehabilitation of aphasic patients who present with significant non-fluency or severely restricted verbal output (Helm-Estabrooks & Albert, 1991). Following the notable effect of the method in three severe non-fluent aphasic subjects (Albert, Sparks, & Helm, 1973) and the report of its efficacy in eight patients in 1974 (Sparks, Helm, & Albert, 1974), several investigators reported significant improvement of their patients after MIT treatment (Popovici, Mihailescu, & Voinescu, 1992).

MIT is a hierarchically structured programme divided into three linguistic levels. In the first two levels, multisyllabic words and short, high-probability phrases are intoned musically and are sung with the patient and tapped out syllable by syllable (Helm-Estabrooks, Nicholas, & Morgan, 1989). The third level introduces longer or more phonologically complex phrases. These longer phrases are first intoned, then produced with exaggerated speech prosody, and finally spoken normally.

As a hierarchically structured programme with a scoring system for each step and well-defined criteria for entering higher levels, MIT is regarded by the American Academy of Neurology (1994) to be among the few language therapy programmes sufficiently precise to allow evaluation across different situations. Objective evidence in terms of efficacy of the method was provided by Belin et al. (1996), who showed the mechanism of recovery from non-fluent aphasia in patients rehabilitated with MIT. They believed that in their patients, the reactivation of Broca's area and the left prefrontal cortex, and deactivation of the counterpart of Wernicke's area in the right hemisphere were the possible mechanisms for recovery from non-fluent aphasia after MIT.

MIT is primarily a method for the rehabilitation of propositional language as reported by Sparks et al. (1974) and others who have replicated its efficacy. Non-fluent aphasic patients with reasonably good auditory comprehension and poor repetition are the best candidates for this method. Transcortical and global aphasic patients and those with evidence of significant posterior language area involvement seem to benefit less from the effects of the method on expressive skills (American Academy of Neurology, 1994). However a few investigators have reported global aphasic patients who showed remarkable progress after MIT therapy (Belin et al., 1996; Van der Lugt-van Wiechen, & Visch-Brink, 1989).

There are a few non-replicated reports regarding side-effects of the method on auditory comprehension skills. Whereas some authors reported ineffectiveness of the method for Wernicke's aphasia (a case report by Helm, 1981) and global aphasia (American Academy of Neurology, 1994) a few investigators showed, as a side-effect, its positive effects in the rehabilitation of the verbal decoding of their patients; for example global aphasic patients reported in the study by Belin et al. (1996), and non-fluent patients reported by Helm-Estabrooks (1983) had better performance in auditory comprehension subtests after MIT therapy. Popovici (1995) applied the method in combination with the method of semantic fields, used it directly for the rehabilitation of auditory comprehension of her patients, and reported good response in Wernicke's aphasic patients.

Although MIT was developed for English-speaking aphasic patients, it has been adapted to some other languages. The purpose of this study was to examine the effect of MIT for Persian aphasic patients. Review of the literature yielded no data concerning this effect in Persian-speaking subjects. As there are prosodic differences between the Persian and English languages, using the method in Persian required adaptations to the original method. The importance of differences in prosody is well described by Seki and Sugishita

(1983), who compared Japanese and English prosodies and adapted the method for the Japanese language. Their comparison showed differences in the use of pitch (Japanese, unlike English, is a tone language) and rhythm (syllable-timed rhythm in Japanese vs stress-timed rhythm in English).

Before the details of this experiment are presented, we give an overview of the prosodic aspects of the Persian language that should be considered when using the MIT method. These aspects that are exaggerated during the method are intonation, points of stress, and rhythm of the stimulus items (Sparks & Deck, 1994).

Intonation

Intonation refers to communication changes in fundamental frequency or the linguistic use of pitch (Hargrove & McGarr, 1994). Persian, like English, and unlike languages such as Chinese, is an intonational language (Haghshenaas, 1990). Intonational languages use intonation for discourse management, for expression of intentions, to specify grammatical function of sentences, and for speech registers. (Speech registers involve the use of special speech patterns in specific contexts or with specific groups of people. Child Directed Speech is one example of a register.) Unlike tone languages such as Chinese, most intonational languages rarely use intonation to differentiate lexical items (Hargrove & McGarr, 1994). The use of intonation for the aforementioned purposes is more exaggerated and pronounced in some Persian accents, for example the accent of people in the city of Esfahan (Sepantaa, 1996).

Stress

With regard to the stress component, however, the characteristics of the English and Persian languages are not the same. Considering the acoustical elements of "lexical stress" (intensity, fundamental pitch, and duration) it has been found that in English syllables, increased duration is the most important factor in perception of stress (Laughlin, Naeser, & Gordon, 1979). The most important factor contributing to stressed syllables in Persian however, is frequency. Sepantaa (1972) showed that stress points in Persian, have frequencies 40–50 Hz higher than the unstressed elements. Khaanlari (1978) found this difference to be five to nine semitones. Khaanlari also had shown earlier that in the Persian language, intonation is a function of so-called "musical accents". In other words, intonational patterns in addition to specification of the above-mentioned characteristics follow the ups and downs of stress points in a sentence.

Rhythm

The Persian language is said to have a "syllable-timed" rhythm. In Persian there are six different, unchangeable syllable patterns (Appendix 1). In other words in Persian, the syllabic pattern CV(C(C)), whether stressed or unstressed, may be uttered with six durations, depending on the vowel type (short or long) and appearance of the last consonants (Haghshenaas, 1990). This feature is different from the English stress-timed rhythm.

Based on the above-mentioned properties of Persian prosody (more important role of pitch in utterance, and using different constant durations), in the present paper the process of adaptation of the MIT technique (the 1989 version by Helm-Estabrooks, Nicholas, and Morgan) for Persian aphasic patients is described in detail. A preliminary report of the effects of the adapted MIT on a number of our Persian patients was published in advance of this paper (Bonakdarpour, Eftekharzadeh, & Ashayeri, 2000). This paper presents the results of the extension of the previous report. The aim of the experiment was primarily to investigate the effects of 15 sessions of MIT treatment in selected patients with non-fluent aphasia; primarily on expository speech (phrase length and number of correct content units) and oral expression skills (repetition, responsive naming, and confrontational naming), and secondarily, on auditory comprehension abilities (word discrimination, commands).

METHOD

Participants

Seven Persian-speaking patients (four males and three females) with non-fluent aphasia served as participants. The Farsi Aphasia Test, or FAT (Nilipour, 1993) and brain CT scans (see later) were used for qualification of patients. Mean age for the patients was 52.4 years (range 45–61 years) and at the time of first evaluation, they ranged from 14 to 58 months (mean = 35.43 months) post-onset. The aetiology of aphasia, as determined by history, physical examination, and brain CT scans, was vascular in origin (due to acute left MCA occlusion in five patients and acute intracerebral haemorrhage in the other two). Therefore five patients appeared to have Broca's aphasia and two had subcortical aphasia (Naeser et al., 1982). The patients had received conventional speech therapy treatment months before MIT was administered. The therapy programme consisted of the Schuell stimulation approach, but unfortunately records of their progress were not available for all patients. As noted by their therapists, their condition and especially verbal output had not changed much after the initial therapeutic sessions.

Based on clinical judgement, the patients were motivated, attentive, and emotionally stable (e.g., they did not have pseudobulbar affect). These factors allowed them to participate in a treatment session. A summary of general characteristics of the patients is represented in Table 1.

Language testing

The FAT is a newly introduced aphasia battery for Persian (Farsi) aphasia and was the only published battery available at the time of our study in Iran. It consists of five sections for the evaluation of aphasia including spontaneous speech, auditory

TABLE 1
General characteristics of patients

	KA	*MRQ*	*MS*	*MD*	*MEM*	*IM*	*FT*
Age	60	54	55	59	47	61	45
Sex	F	M	F	M	F	M	M
MPO	28	14	27	33	57	54	35
Aetiology	CVA	CVA	CVA	CVA	CVA	CVA	CVA
Lesion	B	SC	SC	B	B	B	B
Prior therapy	+	+	+	+	+	+	+

General characteristics of patients selected for the MIT programme with regard to age, sex, months post-onset of acquiring aphasia (MPO), history of prior therapy, aetiology, and site of brain lesion. CVA: Cerebrovascular accident. Type of brain lesions, according to Naeser and Hayward system (1978), are represented by following abbreviations. B: Broca, SC: Subcortical.

comprehension, oral expression, reading, and writing, and is similar to the Boston Diagnostic Aphasia Examination in many aspects. A summarised description of the FAT subtests used for qualification of subjects can be found in Appendix 3.

Using the section in the FAT for the evaluation of spontaneous speech the patients were asked to tell the story of a presented picture. Their outputs were then recorded and analysed based on narrative analysis used by Nilipour (1992). All subjects had speech output with phrase length of two words or less (mean = 1.43) and spontaneous output with less than 20 (mean = 8.3) words per minute that placed them in the non-fluent aphasia category. They all had moderate to good auditory comprehension (above 50% on the FAT, mean = 74.36) and moderately to severely impaired repetition (repetition for high-probable phrases equal to or less than three out of five, mean = 1.71). Table 2 summarises characteristics of the patients based on language testing.

In addition to the above-mentioned variables, confrontational naming, responsive naming (from the oral expression section of the FAT), word discrimination, and commands (from the auditory comprehension section of the FAT) were evaluated based on variables selected in prior studies. To evaluate patients' descriptive speech we also used the variable ''number of correct content units'' (NCCU). NCCU was adapted from an outcome measure named the Index of Lexical Efficacy (ILE), which measures the ratio of informative words to total words produced by the patient. Menn, Helm-Estabrooks, and Ramsberger (1994) first introduced this index as a part of the Communicative Effectiveness Profile (CEP), which is an approach to analysing narrative pictures. Parallel performance between narrative description of the Boston Diagnostic Aphasia Examination (BDAE) ''Cookie Theft'' picture and conversational skills has been shown by Easterbrooks, Brown, and Perera (1982, pp. 93–107). As the mean for the ILE ratio has not been defined in normal Persian-speaking persons we only used the variable NCCU for analysing patients' performance in the picture description task of FAT.

NCCU is the number of discrete units or new bits of information supplied by the patient counting only the ones that describe correct elements of the picture such as *the man*, *is reaching*, and *for a bird* (in the FAT pictures). There are 25 bits of information in the pictures and the participants of this study could not produce more than 13 items out of 25.

With regard to auditory comprehension other authors have also measured complex ideational material (CIM) as a subtest. However this variable was left out of our study. We believed that CIM could not be appropriately measured by the FAT due to cultural bias resulting from translation of this section from the BDAE, and lack of accuracy in devising the controlling questions during adaptation from the BDAE.

TABLE 2
Characteristics based on language testing

	KA	*MRQ*	*MS*	*MD*	*MEM*	*IM*	*FT*
Speech output							
Phrase length (words)	0	0	2	2	2	2	2
Word per minute	0	3	4.33	12	17.53	14	7.25
Auditory comprehension (%)	69.6	55.3	57.2	68	94	78.4	98
Repetition (high-probability)	0	2	0	1	3	3	3

Characteristics of patients in terms of spontaneous speech output, auditory comprehension, and repetition of high-probability phrases of the Farsi Aphasia Test.

In addition to target variables a number of other variables, presumed not to be affected by MIT treatment, were measured as baselines controlling target outcome measures. These variables were selected from the reading and writing sections of the FAT and included symbol and word recognition, word identification, words to dictation, and sentences to dictation.

CT scan localisation of brain lesions

To localise brain lesions on CT scans, Naeser and Hayward's labelling system (1978) was employed. The Naeser and Hayward labelling system includes six slices, each representing 1cm thickness of the head, that are named according to the known cortical language areas present in each. The slices, from the bottom to top, contain Broca's area (B), Broca's and Wernicke's area (B/W), Wernicke's area (W), and supramarginal and angular gyrus (SM); SM+1 and SM+2 slices are one centimetre and two centimetres above the SM slice respectively. Using this system Naeser and Hayward described the pattern of brain involvement in major cortical aphasia syndromes (i.e., Broca's, transcortical motor, Wernicke's, global, and conduction aphasia) and Alexander, Naeser, and Palumbo (1987) described patterns of subcortical syndromes (i.e., capsular/putaminal aphasias).

Brain lesions in all patients—except one (MD) with patchy involvements in Wernicke's area—matched the criteria suggested by Naeser and Helm-Estabrooks (1985). According to Naeser and Helm-Estabrooks good candidates for the MIT programme do not have a lesion in the temporal isthmus at slice B/W, a large lesion in the Wernicke's area at slice W, and any lesion in the right hemisphere. The composite CT scan lesion sites for the patients are shown in Figures 1 and 2.

Experimental design

Using a pre, post treatment design the subjects were tested twice (at 1-month interval) before MIT treatment was administered (First treatment-free phase). After the first treatment-free phase MIT was administered at the rate of 3–4 days per week and for 1 month (Treatment phase). Subjects were subsequently tested following completion of the treatment period, and a month later as follow-up. (Second treatment-free phase)

Pre-treatment phase

To control for spontaneous recovery, at each assessment during this phase, narrative and spontaneous speech samples were obtained and analysed with regard to phrase length and NCCUs. Also at each testing session oral expression variables (i.e., repetition, responsive naming, and confrontational naming) and auditory comprehension subtests (i.e., word discrimination and commands) were measured using the FAT. Hence seven target and four control variables were evaluated at each assessment.

MIT phase

The MIT phase included stimulus items that were intoned and recited with the patients. The treatment provided consisted of three levels (see Appendix 4).

Stimulus items. Items used during the MIT treatment phase included high-probability words, phrases, and sentences accompanied by adapted pictures or environmental clues. In each session 15–20 stimuli were practised. The stimuli were not the same in

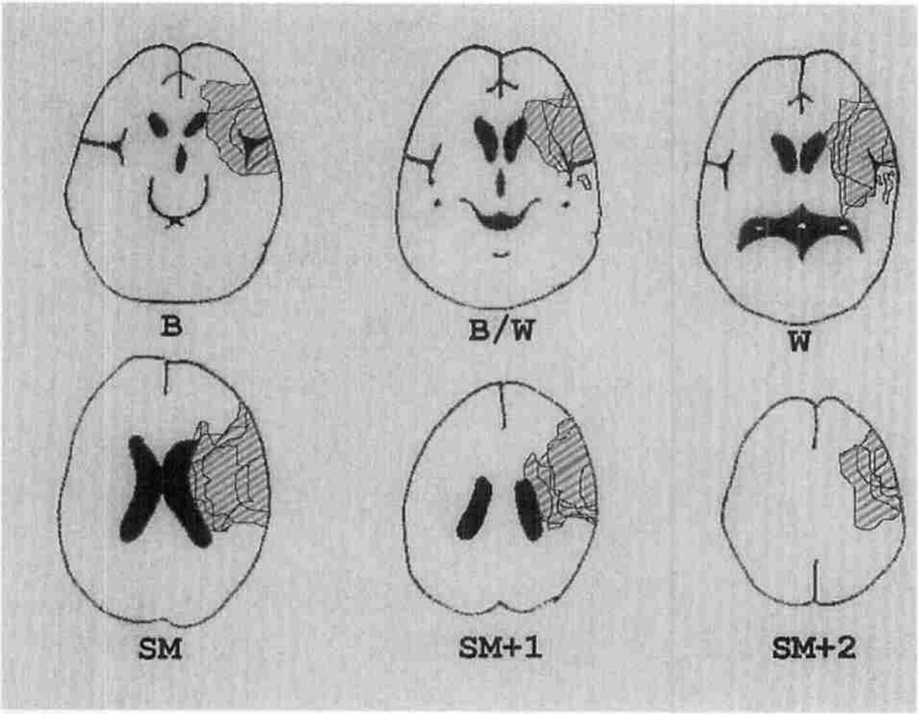


Figure 1. Composite CT scan for the five right-handed patients with Broca's aphasia, using the Naeser and Hayward labelling system (1978).

consecutive sessions and were rotated through a series of treatment sessions. Items were chosen according to each patient's communication needs as reported by their caregivers and close relatives.

Target items were also selected based on their phonological difficulty (Hatfield, 1972; Helm-Estabrooks & Albert, 1991), syntactic complexity as per the Helm elicited programme for syntax stimulation (Helm-Estabrooks & Ramsberger, 1986), and number of syllables. As to the phonological difficulty, we first started with labial phonemes at the first level and proceeded to coronal and dorsal phonemes at the later levels. The syntactic hierarchy used in this study is depicted in Appendix 2. These characteristics will be noted later under descriptions of each level.

Intoning stimulus items. To intone the stimuli, i.e., to exaggerate the prosodic pattern, three prosodic components of pitch, points of stress, and rhythm are altered. As mentioned previously, because intonation and stress patterns are interrelated in Persian, it is sufficient to deal with pitch and duration for intoning Persian stimuli.

In melodic intonation varying pitch of speech is reduced and stylised into melodic pattern involving the constant pitch of intoned notes (Figure 3). Melodic intonation uses a limited range of musical notes. As reported by the American authors, the musical pattern of melodic intonation in the first and second levels is similar to that of recitative style in operas (Sparks & Deck, 1994; Sparks & Holland, 1976). As used in MIT, in this style, unlike songs, a limited range of musical notes is used (three to four whole tones in the original method). In this regard syllables are uttered using several musical notes at the same pitch (one note for each syllable) and a higher/lower musical note for a higher/lower

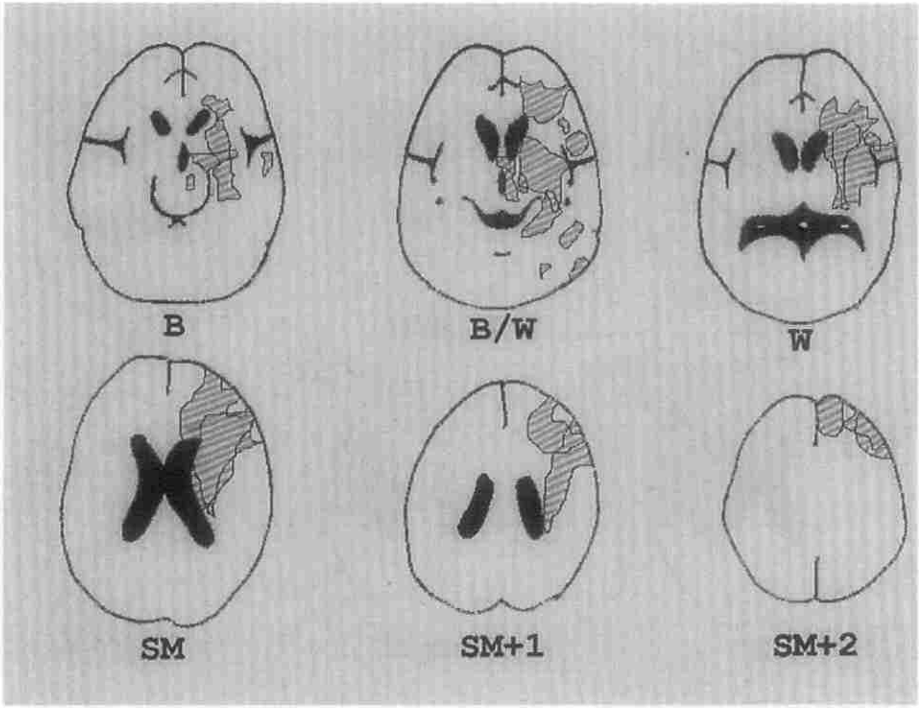


Figure 2. Composite CT scan for the two right-handed patients with subcortical capsular/putaminal mixed aphasia, using the Naeser and Hayward labelling system (1978).

syllable pitch in the related word or sentence. In this manner the pattern of spoken prosody is transposed to a musical melody (see Figure 3).

The Mahour scale of Persian music, which is the closest scale to the Western major scale, was used for intoning the stimuli. The counterpart of the recitative style in Persian is the Gousheh pattern. In Persian music (like modal Western music) not all the pitches in a scale (like a major scale) are used for composing a song section; rather, a limited number of four ascending or five descending notes are applied for this purpose. So given the Mahour scale with eight pitches the ascending and descending frames are as shown in Figures 4 and 5. These patterns of asciending or descending frames are called Gousheh. A very helpful vocal example in this regard is the song ''Veil of the soul's visage'' by Hassan Riahi. On the basis of what previous research on patterns of stress in Persian had revealed we could use a range of 2.5 to 4.5 whole tones while intoning the stimulus items. The broader range of pitches was

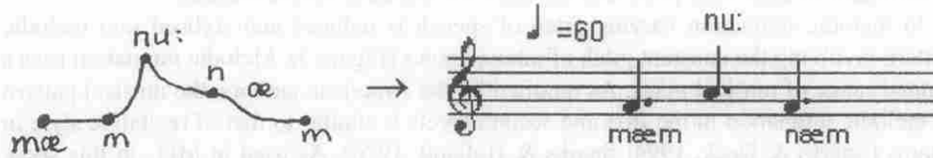


Figure 3. Transposition of varying speech pitch into melodic pattern in MIT: The figure on the left shows the varying pitches in the Persian phrase ''mamnoonam'' (thanks). In the MIT method the phrase is intoned using notes with constant pitches. (Right) ♩ = 60 indicates that there should be 60 crotchets in a minute, so ♩ and ♩. stand for 1 and 1.5 second durations respectively.

Figure 4. The Mahour scale.

Figure 5. Ascending and descending frames in the Mahour scale.

used when the sentences were more complex (specially near the end of level two and at level three).

With regard to duration of the intoned syllables, as mentioned previously, there are six syllable durations in Persian (Appendix 1). We used three durations for intoning the items, i.e., short, medium, and long. These durations were transposed to notes with durations of 1, 1.5, and 2 seconds respectively. Items should be performed with a slow tempo. By using these duration patterns the syllable-timed rhythm of the phrases shapes up. It is noteworthy that without using the different durations the resultant intoned word or phrase may seem strange or humorous. Here what is important is the proportion of the durations, and relative freedom can be used.

Melodic intonation hierarchy and linguistic content of MIT levels. A brief description of methodology of each level including changes made for adapting the method for Persian follows (see Appendix 4 for a summary of each level). For more detailed information regarding the methodology introduced by the American authors, the reader is referred to Helm-Estabrooks et al. (1989).

Level I: This level consisted of five steps. It began with humming the stimulus item and went through unison singing of the clinician and patient, unison singing with fading of the clinician halfway through the singing of the item, immediate repetition of the target item by the patient after the clinician, and response to a probe question asked by the clinician regarding the item. The stimulus items in the first level included the simplest forms with regard to number of syllables, phonological difficulty, and syntactic complexity (see Appendix 5 for examples). The target items consisted of two or three syllables. They included bilabials and visualisable sounds like /p/ and /l/ and consisted of simple nouns and the easiest syntactic structures, such as imperative and interrogative sentences as per the Helm elicited programme for syntax stimulation or HELPSS. As the stimulus items consisted of two or three syllables in the first level, a simple pattern of high and low pitches, with limited variation in pitch, was employed for intoning (Figure 6 depicts an example for this level).

Level II: This level used more complex words and sentences with regard to number of syllables, phonology, and grammatical structure. At this level coronal and dorsal

Figure 6. Example of an intoned phrase in level I. Note the aspects of intoning including pitches and duration of syllables. Two pitches are used, the interval between which is not more than two tones. The duration for each of the syllables is 1.5 second.

phonemes were introduced based on the abilities of each patient. Also more complex syntaxes could be administered.

Level II had four steps (introduction of the intoned item by the clinician, unison singing with fading, delayed repetition of the target item by the patient, and response to probe question) and introduced ''delays'' between stimulus and response. If a patient was unable to complete a step with delays, he or she was permitted to back up to the previous step (see Appendix 5).

Laughlin et al. (1979) suggested increasing the duration of syllables when the patient had problem in production of the stimulus, and this may facilitate the progress of the patient through the steps. As pauses can improve the comprehension of the stimuli (Darley, 1982), when possible we also made use of embedded pauses with at least 1-second duration between words of a stimulus (see Figure 7). The process of intoning at this level was similar to the first level but more complex use of pitch was used as the number of syllables increased (see Figure 7).

Level III: The final level of MIT is designed to return the speech of the patient to normal prosody. This was achieved through a transitional technique called ''sprechgesang'' or ''speech song''. In this technique, rhythm and stress of each target phrase were accentuated while the intonational characteristics used in previous levels were dropped and replaced by the constantly changing pitch of normal speech (Helm-Estabrooks & Albert, 1991). ''Sprechgesang'' is used in choral speaking but Schoenberg has used it more lyrically in his ''Ode to Napoleon'' and ''Pierrot Lunaire'' (Sparks & Deck, 1994). In Persian this type of speaking can be found in ceremonial choral speaking, in traditional narrations (Naghghaali), and more lyrically by some musicians (e.g., Alireza Assar). In level III, target stimuli were much more difficult than the previous levels in terms of number of syllables, phonology, and syntactic structure (see Appendix 5).

Monitoring of progress as a result of therapy. A scoring system was used for each step of each MIT level, allowing progress (or lack of progress) of the therapy to be

Figure 7. Example of intoned phrase in level II or III. Note wider range of pitches, more variable pitches, and use of a pause between the two verbs [''Mikhaastam'' and ''biyaay'' (I wanted you to come)].

charted. If the patient did not produce the stimulus item in each step of the first level, that item was not scored and the next item introduced. Patients were allowed to progress from one level of MIT to the next if they had earned an overall score of 90% or better for five consecutive sessions. At the second and third levels the patient was permitted to back up to the previous step, but this would lower their total score during each session and slowed moving from one level to the other.

In the present study 15 sessions of treatment were administered, this decision was based on observation that the first signs of improvement in target variables might be observed after 15 sessions of treatment (Helm-Estabrooks et al., 1989).

Post-treatment phase

All variables measured during the pretreatment phase were measured again after completion of treatment in order to examine the effect of treatment after one month.

Reliability

Data collection at each phase was done independently by two individuals. Disagreements were resolved by a third scorer to result in near 100% agreement.

RESULTS

Our observations during the treatment sessions showed that, as expected, all patients seemed to have better output when the items were intoned. When the patients had problems producing the intoned stimulus, increasing the duration of syllables (but preserving the rhythmical pattern) helped to improve their output. Lengthening the pauses between the words of the target phrase to at least 1 second also helped.

Statistical analysis, using a pre and post Wilcoxon signed rank test during the two treatment-free phases and during the treatment phase, was done to test the significance of changes in each of the seven target variables and the four control variables. Each variable was tested separately from the others.

Analysis of the data showed no significant progress in any of the aforementioned variables during the two MIT-free phases (Tables 3 and 4). Also no significant deterioration was seen throughout the two treatment-free phases. Conversely, comparison between pre- and post-MIT using a Wilcoxon signed rank test showed statistically significant improvement (Table 3 and Figure 8) in subtest scores (phrase length: p = .0125; number of correct content units: p = .0107; confrontational naming: p = .0312; responsive naming: p = .0107; repetition: p = .0084; word discrimination: p = .0238; commands: p = .0238).

As shown in Table 4, pre- and post-treatment assessment of non-target variables, i.e., reading and writing subtests (spelling to dictation, sentences to dictation, word identification, and symbol and word recognition), after MIT therapy showed no significant improvement.

DISCUSSION

The results of this study showed statistically significant improvement in target variables (i.e., phrase length, NCCU, repetition, confrontational naming, responsive naming, word discrimination, and commands) using MIT. Non-target variables (spelling to dictation, sentences to dictation, word identification, and symbol and word recognition) did not improve during the MIT and post-MIT phases after 15 sessions of treatment. So it seems

TABLE 3
Comparing probability levels for three phases

	$T_1 - T_2 <> 0$	$T_2 - T_3 < 0$	$T_3 - T_4 <> 0$
Variables	*p*	*p*	*p*
Confrontational naming	0.54	0.0312	0.79
Responsive naming	0.21	0.0107	0.66
Repetition	0.21	0.0084	0.60
Word discrimination	0.91	0.0238	0.91
Commands	0.56	0.0238	0.49
Phrase length	1.00	0.0125	0.13
NCCU	0.49	0.0107	0.44

T_1: Score of the first assessment (before treatment); T_2: Score of the second assessment (at the beginning of the treatment period); T_3: Score of the third assessment (at the end of the treatment period); T_4: Score of the fourth assessment (one month thereafter); *p*: Probability level; NCCU: Number of Correct Content Units.

TABLE 4
Probability levels for the untreated baseline variables in the treatment phase and the two treatment-free phases

	$T_1 - T_2 <> 0$	$T_2 - T_3 < 0$	$T_3 - T_4 <> 0$
Variables	*p*	*p*	*p*
Symbol & word discrimination	0.49	0.1	1
Word recognition	0.28	0.49	0.42
Sentences to dictation	1	1	1
Spelling to dictation	0.16	0.49	1

For each variable there has been no significant change in each phase.

that therapy with MIT was efficacious in the rehabilitation of the target variables after 15 sessions of treatment.

The statistically significant results obtained during the treatment phase cannot be accounted for by spontaneous recovery during a short period of time (4–5 weeks) because the MIT programme was instituted after more than 14 months post-onset and there was no significant improvement during the non-treatment phases. Furthermore it might be less probable that the improved results would be non-specific consequences of the intervention or due to placebo effect or chance. The improved findings are presumed to be a result of generalisation from what was treated to what was tested; only the selective target variables (and not the non-target variables) were improved.

Effects of MIT on verbal output of the patients

The primary goal of the MIT technique, as originally described by Sparks et al. (1974), is the rehabilitation of propositional speech. Statistical analysis showed that in our patients the NCCU and phrase length increased significantly during the treatment phase. At the same time there were no significant changes in the two treatment-free periods and in the four baseline variables. Notably, the increased performance of the patients remained

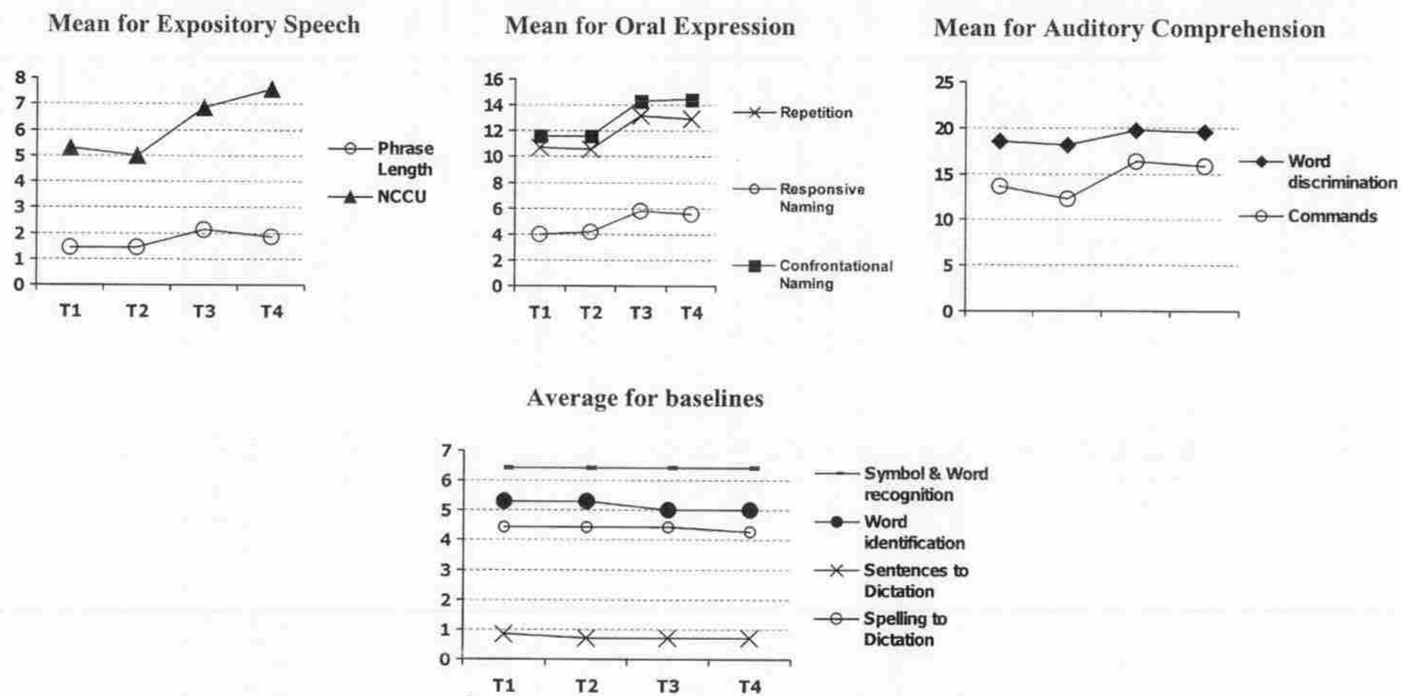


Figure 8. The three upper charts show patients' improvement in expository speech, oral expression, and auditory comprehension variables after treatment with MIT, while there has been no significant improvement or deterioration in the two treatment-free phases (see Table 3). At the same time there have been no changes with regard to the baseline (untreated) variables (see Table 4). NCCU = Number of correct content units.

constant in the second treatment-free phase. In this regard, one might infer that the observed changes could be the result of MIT treatment sessions. The significant improvement in phrase length is in agreement with the findings reported by Sparks et al. (1974).

We presume that the patients' statistically significant progress in the outcome measure NCCU would correlate with better performance in actual conversational skills at home or in society, as relatives and caregivers of the patients also reported. But as there are no objective data regarding the association between improvement of NCCU and improvement of communicative skills in the Persian language, strong conclusions cannot be drawn before evidence is provided by future studies.

Further evidence regarding the possible effectiveness of MIT in our patients was their significant improvement in oral expression subtests of the FAT. It is believed that the process of improvement first begins in progress in the oral expression subtests and then extends to expository speech. However, this is not a goal for therapy, and the mentioned variables might improve while no change might be present in patients' conversational ability, as Helm-Estabrooks and colleagues reported in their cases, described as poor candidates for MIT (1989). Our results are in agreement with findings reported by Sparks et al. (1974) and those observed by French (Belin et al., 1996) and Romanian (Popovici et al., 1992) contributors.

There is now objective evidence that shows how MIT works. In a PET study, Belin et al. (1996) showed that MIT reactivated essential motor language zones, such as Broca's area and prefrontal cortex in the left hemisphere, while reducing the abnormal activation in the right hemisphere. The role of increased duration for MIT efficiency during treatment was shown by Laughlin et al. (1979). Also there is evidence that Broca's area plays a role in rhythm processing (Fiez, Raichle, Tallal, & Petersen, 1993). Whether exaggerated pitch patterns would have the same effect is yet to be investigated by future studies.

Effects of MIT on auditory comprehension

A few authors have reported, as a side-effect, improvements in auditory comprehension abilities of their patients (Belin et al., 1996; Helm-Estabrooks, 1983; Popovici, 1995), notably in the variables word discrimination, commands, and complex ideational material from the BDAE. In our series of patients, improvements in word discrimination and auditory commands turned out to be statistically significant and were constant after at least a 1-month follow-up.

Although Sparks et al. did not find changes in auditory commands, complex ideational material, and reading comprehension to be significant, Helm-Estabrooks (1983) reported significant improvement in complex ideational material in a group of 22 non-fluent aphasics treated with MIT. Naeser and Helm-Estabrooks (1985) also reported significantly increased auditory comprehension in a number of their cases. French authors (Belin et al., 1996) also revealed improvements in verbal discrimination, orders execution (auditory commands), and logic and reasoning (complex ideational material). Interestingly, through a controlled trial, Popovici (1995) showed MIT to be effective on verbal decoding, particularly in Wernicke's and to a lesser extent in Broca's and anomic aphasic patients.

Three factors in the intoned stimuli could affect auditory comprehension: altered duration of the syllables, introducing pauses between the words, and transforming the varying pitch of normal speech to constant pitch of musical notes. Intoned phrase items with increased inter-syllabic pause time and slowed rates of presentation enhance auditory comprehension in aphasia (Helm-Estabrooks, 1983) The role of slowing the overall rate of speech in improvement of auditory comprehension is quite well known (Albert & Bear, 1974; Gardner, Albert, & Weintraub, 1975; Lasky et al., 1976; Weidner & Lasky, 1976). This is true because increasing the syllable duration and using pauses between words provides extra processing time for the patient. At the same time those changes enhance the language signal, as exemplar phonetic prototypes of the words are produced in supranormal intonation contours. Helm-Estabrooks reported a patient who could not comprehend a word when produced with normal prosody, but was able to understand it when the word was intoned (Helm,1981).

Although the aforementioned studies support the effect of prosodic changes in enhancing auditory comprehension, this effect is not large and so it would be better to regard it as a side-effect of MIT treatment. Also not all aphasic patients show the same amount of improvement. The effect is greatest in the most severely impaired aphasic patients in terms of auditory comprehension, e.g., Wernicke's aphasic patients (Blumstein, Katz, Goodglass, Shrier, & Dworetsky, 1985; Kimelman, 1999; Popovici, 1995), and least or absent in Broca's aphasic patients. It is yet to be determined whether the melodic component of MIT affects auditory comprehension.

Considerations regarding duration of treatment

Most previous studies reviewed in the introduction of this paper have not indicated the duration of treatment periods, but rather, completion of the MIT programme has been considered. In a preliminary report on the effectiveness of MIT, Albert et al. (1973) described three patients who made gains after about 2–6 weeks of MIT treatment. In another review based on the Naeser and Helm-Estabrooks study, about 120 sessions were indicated to produce considerable results in two non-fluent patients (Helm-Estabrooks et al., 1989). The assessment report from the American Academy of Neurology (1994) has described MIT to be best given in short, frequent sessions during a limited span of time, namely 3–6 weeks. Helm-Estabrooks and colleagues (1989) recommended a minimum of 15 sessions of treatment before the first assessment of the patients. Our patients received 15 sessions of treatment during 1 month and showed some improvement. This study, therefore, provides evidence in favour of the suggestion that improvement can be detected after 15 sessions. There is still no objective evidence regarding the fact that progress within this period can predict further improvement if the treatment is continued.

Points for maximising progress of the patients during a session

As mentioned in the methodology, patients may have problems producing the stimulus items during an MIT treatment session. This results in lower scores during each MIT session and a delay in moving from one level to the other, and therefore the whole treatment process slows down. However, as observed during MIT treatment sessions, by using certain methods, production of stimuli could be facilitated.

There are few studies concentrating on maximising patients' output during an MIT session. One example is the study by Laughlin et al. (1979) who showed that with prolongation of syllables patients could pass MIT steps more easily. We observed the same result, but given that stress points in Persian are more related to the increase of the pitch rather than duration, we presume that exaggerating the inflection pattern might also be of great importance in facilitating patients' output. Validity of this statement has yet to be examined however, especially when considering the significance of the effect of syllable duration according to Belin et al. (1996).

Another factor that seemed to facilitate the patients' output was lengthening of pauses between the words within a phrase. Pauses provide speakers the opportunity to process or plan abstract linguistic information (Darley, 1982). Moreover, considering that defects in auditory comprehension are usually present in aphasic patients, placing intra-turn pauses may possibly aid comprehension and thus maximise output. This can be taken into account when using the MIT technique.

Our study has obvious limitations for decisive conclusions. FAT is a newly developed battery and there was no research available concerning the reliability and validity of it at the time of our study. Also, stronger conclusions could be drawn if there had been a control group and if the patients had received the treatment for more than 15 sessions. Both of these were not feasible because of lack of support for the study. Finally, although we presumed that the syntactic complexity in English and Persian are similar, objective data are still necessary to prove this. Studies on the syntactic properties of Persian aphasia and the classification of different syntactic patterns are currently lacking.

CONCLUSION

Our study showed that MIT can be adapted for Persian aphasia and administered with measurable positive results after 15 sessions of treatment. MIT improved primarily spontaneous speech production and, as minor effects, selected oral expression and auditory comprehension subtests. This can be regarded as a short-term effect in agreement with other investigations that have reported efficacy of MIT in other languages. Still more long-term follow-up study with randomised controlled clinical trials is needed for stronger conclusions.

Manuscript received 24 April 2001
Manuscript accepted 8 April 2002

REFERENCES

Albert, M., & Bear, D. (1974). Time to understand. A case study of word deafness with reference to the role of time in auditory comprehension. *Brain, 97*, 373–384.

Albert, M.L., Sparks, R., & Helm, N. (1973). Melodic intonation therapy for aphasia. *Archives of Neurology, 29*, 130–131.

Alexander, M.P., Naeser, M.A., & Palumbo, C.L. (1987). Correlation of subcortical CT lesion and aphasia Profiles. *Brain, 110*, 961–991.

American Academy of Neurology (1994). Assessment: Melodic intonation therapy [Report of the therapeutics and technology assessment subcommittee of the American academy of neurology]. *Neurology, 44*, 566–568.

Belin, P., Van Eeckhout, P., Zibovicius, R.P., Francois, C., Guillaume, S., Chain, F., Rancurel, G., & Samson, Y. (1996). Recovery from nonfluent aphasia after melodic intonation therapy: A PET study. *Neurology, 47*, 1504–1511.

Blumstein, S., Katz, B., Goodglass, H., Shrier, R., & Dworetsky, B. (1985). The effects of slowed speech on auditory comprehension in aphasia, *Brain and Language, 24*(2), 246–265.

Bonakdarpour, B., Eftekharzadeh, A., & Ashayeri, H. (2000). A preliminary report on the effects of Melodic Intonation Therapy in the rehabilitation of Persian aphasic patients. *Iranian Journal of Medical Sciences, 25*(3&4), 156–160.

Darley, F.L. (1982). *Aphasia* (pp. 217–218). Philadelphia: WB Saunders Company.

Easterbrooks, A., Brown, B.B., & Perera, K. (1982). A comparison of the speech of adult aphasic subjects in spontaneous and structured interactions. *British Journal of Disorders of Communication, 17*(3), 93–107.

Fiez, J.A., Raichle, M.E., Tallal, P.A., & Petersen, S.E. (1993). Activities of a left frontal area near Broca's area during auditory detection and phonological access task [abstract]. *Journal of Cerebral Blood Flow and Metabolism, 13*(suppl 1), S504.

Gardner, H., Albert, M.L., & Weintraub, S. (1975). Comprehending a word: The influence of speed and redundancy on auditory comprehension in aphasia. *Cortex*, *11*(2), 155–162.

Haghshenaas, A.M. (1990). *Phonetics* (pp. 127–129). Tehran: Aagaah Publisher. [In Persian]

Hargrove, P.M., & McGarr, N.S. (1994). *Prosody management of communicative disorders* (p. 153). London: Whurr Publishers.

Hatfield, F.M. (1972). Looking for help from linguists. *British Journal of Disorders of Communication*, *7*(1), 64–81.

Helm, N. (1981). Treatment of comprehension disorders. In M. Albert, H. Goodglass, N. Helm, A. Rubens, & M. Alexander (Eds.), *Clinical aspects of dysphasia* (p. 161). New York: Springer-Verlag.

Helm-Estabrooks, N. (1983). Exploiting the right hemisphere for language rehabilitation: Melodic Intonation Therapy. In E. Perecman (Ed.), *Cognitive processing in the right hemisphere* (pp. 229–240). New York: Academic Press.

Helm-Estabrooks, N., & Albert, M.L. (1991). *Manual of aphasia therapy* (pp. 207–228 & 240). Austin, TX: Pro-Ed.

Helm-Estabrooks, N., Nicholas, M., & Morgan, A.R. (1989). *Melodic intonation therapy kit*. Chicago: Applied Symbolix.

Helm-Estabrooks, N., & Ramsberger, G. (1986). Treatment of agrammatism in long-term Broca's aphasia. *British Journal of Disorders of Communication*, *21*, 39–45.

Khaanlari, P.N. (1978). *A history of Persian language (First Volume)* (p. 90). Teheran: The Foundation for Iranian Culture Publishing. [In Persian]

Kimelman, M.D.Z. (1999). Prosody, linguistic demands, and auditory comprehension in aphasia, *Brain and Language*, *69*(2), 212–221.

Lasky, E., Weidner, W., & Johnson, J.P. (1976). Influence of linguistic complexity, rate of presentation and interphrase pause time on auditory-verbal comprehension of adult aphasic patients. *Brain and Language*, *3*(3), 386–395.

Laughlin, S.A., Naeser, M.A., & Gordon, W.P. (1979). Effects of three syllable durations using the melodic intonation technique. *Journal of Speech and Hearing Research*, *22*, 311–320.

Menn, L., Ramsberger, G., & Helm-Estabrooks, N. (1994). A linguistic communication measure for aphasic narratives. *Aphasiology*, *8*(4), 343–359.

Naeser, M., Alexander, M., Helm-Estabrooks, N., Levine, H., Laughlin, S., & Geschwind, N. (1982). Aphasia with predominantly subcortical lesion sites. Description of three capsular/putaminal aphasia syndromes. *Archives of Neurology*, *39*, 2–14.

Naeser, M.A., & Hayward, R.W. (1978). Lesion localization in aphasia with cranial computed tomography and Boston Diagnostic Aphasia Exam. *Neurology*, *28*, 545–551.

Naeser, M.A., & Helm-Estabrooks, N. (1985). CT lesion localization and response to Melodic Intonation Therapy with nonfluent aphasia cases. *Cortex*, *21*, 203–223.

Nilipour, R. (1992). Objective indexes for the evaluation of speech. *Majalle-ye Zabaanshenaasi [Journal of Linguistics]*, *9*(1), 49. [In Persian]

Nilipour, R. (1993). *Farsi Aphasia Test*. Tehran: Iran University of Medical Sciences Press. [In Persian]

Popovici, M. (1995). Melodic intonation therapy in the verbal decoding of aphasics. *Romanian Journal of Neurology and Psychiatry*, *33*(1), 57–97.

Popovici, M., Mihailescu, L., & Voinescu, I. (1992). Melodic Intonation Therapy in the rehabilitation of Romanian aphasics with buccolingual apraxia. *Review of Romanian Neurology and Psychiatry*, *30*, 99–113.

Seki, K., & Sugishita, M. (1983). Japanese-applied Melodic Intonation Therapy for Broca aphasia. *No To Shinkei*, *35*(10), 1031–1037. [Abstract]

Sepantaa, S. (1972). *An electro-acoustical study of Persian phonemes*. Doctoral thesis submitted to the faculty of letters, Tehran University. [Persian abstract]

Sepantaa, S. (1996). Laboratory analysis of Persian accents in few central Iranian regions and Esfahan. *Seasonal Journal of the Academy of Persian Language and Culture*, *2*(2), 82–104. [Persian]

Sparks, R.W., & Deck, J.W. (1994). Melodic intonation therapy. In R.C. Chapey (Ed.), *Language intervention strategies in adult aphasia* (pp. 368–379). Baltimore: Williams & Wilkins.

Sparks, R., Helm, N., & Albert, M.L. (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex*, *10*, 303–336.

Sparks, R., & Holland, A. (1976). Method: Melodic intonation therapy. *Journal of Speech and Hearing Disorders*, *41*, 287–297.

Van der Lugt-van Wiechen, K., & Visch-Brink, E.G. (1989). Die Melodic intonation therapy bei Patienten mit einer globalen aphasie; das hemmen von recurring utterances. *Sprache-Stimme-Gehoer*, *13*, 142–145.

Weidner, W., & Lasky, E. (1976). Interaction of rate and complexity of stimulus on performance of adult aphasic subjects. *Brain and Language*, *3*, 34–70.

APPENDIX 1

Different syllable durations in Persian

Syllabic pattern	*Example*	*English*
cv̆	/be/ in /beʃin/	(Sit)
cv̄	/bʌ/ in /bʌbʌ/	(Dad)
cv̆c	/dær/	(Door)
cv̄c	/ʃir/	(Milk)
cv̆cc	/goft/	(S/he said)
cv̄cc	/rʊxt/ in /fɔrʊxt/	(S/he sold)

APPENDIX 2

Syntactic hierarchy of stimulus items

1. Imperative intransitive
2. Imperative transitive
3. Interrogative
4. Declarative transitive
5. Declarative intransitive
6. Comparative
7. Passive
8. Yes/No question
9. Direct or indirect object
10. Embedded sentences
11. Future

Based on Helm-Estabrooks and Ramsberger, 1981.

APPENDIX 3

Brief description of the FAT variables evaluated in the study

	Description of items tested by the FAT
Spontaneous speech	
Phrase length	Number of words in the longest uninterrupted word runs in spontaneous speech. Normal subjects produce a mean of 6.23 words for each uninterrupted word run (Nilipour, 1992)
Words per minute	Number of words in spontaneous speech produced during description of the pictures in the FAT. Normal subjects produce a mean of 93.26 words/min (Nilipour, 1992)
Auditory comprehension	
Total auditory comprehension score	Contains total of 51 items. The total score is expressed in percentages
Word discrimination	The patient is asked to point to the picture or object after being presented with the word (20 items)
Commands	The patient is asked to carry out commands given by the evaluator (5 items)
Oral expression	
Repetition	Total of 20 items (10 items for word repetition, 5 items for high-probability phrases and 5 items for low-probability phrases)
Confrontational naming	The patient is presented with pictures or real objects and is asked to name them (25 items)
Responsive naming	The patient is presented with a stimulus question and is asked to supply a one-word response to the question (10 items)
Reading	
Basic symbol and word recognition	The patient is asked to match a written case or script with the suitable item in a series of cases or scripts (10 items)
Word identification	The patient is asked to find a word (read by the evaluator) among a series of written words (10 items)
Writing	
Spelling to dictation	The patient is asked to print dictated letters and words (10 items)
Sentences to dictation	The patient is asked to print dictated sentences (5 items)

APPENDIX 4

Overview of melodic intonation therapy

Level	*Step*	*Number of points*
I.	1. Humming	1
	2. Unison singing	1
	3. Unison singing with fading	1
	4. Immediate repetition	1
	5. Response to a probe question	1
II.	1. Introduction of item	no score
	2. Unison with fading	1
	3. Delayed repetition	2
	(Back up: Unison with fading)	1
	4. Response to probe question	2
	(Back up: Delayed repetition)	1
III.	1. Delayed repetition	2
	(Back up: Unison with fading)	1
	2. Introducing sprechgesang	no score
	3. Sprechgesang with fading	2
	Back up: Unison sprechgesang)	1
	4. Spoken repetition with delay	2
	(Back up: Sprechgesang with fading)	1
	5. Response to probe question	2
	(Back up: Delayed repetition)	1

APPENDIX 5

Sample stimulus items for different levels of MIT

Level I	
/bijʌ/	Come
/bɔrɔ/	Go
/bebin/	Look
/dævʌ/	Pills
/ʔʌbi/	Blue
/mæmnʊnæm/	Thank you
/piræнæm/	My shirt
/tʃetori/	How're you?
/behtæræm/	(I'm) better
/tæmʊme/	(It's) finished
Level II	
negʌ kɔn	Look
zænge dær	Doorbell
bɔlæn ʃɔ	Wake up
befærmʌʔin	Come in (please)
sæd tɔmæn	100 touman(s) [money]
ki ʔʊnjʌs	Who's there?
tʃi ʃɔde	What's up?
ʃirɔ bæst	(S/he) turned off the tap
bʌd miyʌd	It's windy
dɔre miz	By (the) table
Level III	
/dæstʃʊʔi/	Bathroom
/kɔdʒʌ miri/	Where're you going?
/bimʌrestʌn/	Hospital
/dævʌmɔ bede/	Give my pills
/pirænɔ bepʊʃæm/	(Should I) put on the shirt?
/biʌ tʊ ʔɔtʌqæm/	Come in my room
/hævʌ gærm ʃɔde/	The weather's got warm
/xʌnʊmeʃæm?	I'm his wife [compact form]
/ʔʊn bɔzɔrgtære/	That (one) is bigger
/sæb kɔn bijʌm/	Wait for me to come

A STUDY OF THE EFFECTIVENESS OF AN ADAPTATION OF MELODIC INTONATION THERAPY IN INCREASING THE COMMUNICATIVE SPEECH OF YOUNG CHILDREN WITH DOWN SYNDROME

Debbie Carroll
Faculty of Music
McGill University
Montreal, Canada.

August 1996

A thesis submitted to the Faculty of
Graduate Studies and Research
in partial fulfilment of the requirements of the degree of
M.A. in Music Education

ABSTRACT

The present study examined the effectiveness of an adaptation of Melodic Intonation Therapy (MIT) in increasing the communicative speech of young children with Down syndrome. Eight children were matched on the basis of their mean length of utterance (MLU) and randomly assigned to one of two groups – spoken or melodic. Children received the same treatment during 12-weekly 30-minute individual sessions, except for the manner in which target phrases were presented: spoken versus melodically intoned. A drum was used with all the children to support the rhythmic patterns of the target phrases.

Scores for three dependent measures - total number of words, mean length of utterance and rate of response (time required to produce 100 consecutive utterances) were obtained from the transcripts of pre- and post-intervention language samples of children at play at home, first with a parent and then with myself. Every verbal response during each weekly session was categorized according to the nature of the response: unison, imitative, conversational and spontaneous. The number of responses within each category and the total number of responses for each session were then computed.

Findings revealed that interconnected contextual factors, such as the physical setting, child-researcher relationship and the play routine influenced verbal output, regardless of group. Specifically, the drum played a key role for all children in increasing the length and clarity of response and was also an important factor in effecting change within the levels of intervention, particularly with regard to the inverse relation between imitative speech and conversational and spontaneous speech.

The only factor that effected a group difference for total verbal output, length of response and rate of response was the melodic versus spoken manner in which the target phrases were presented. A comparison of the pre-and post-intervention scores for the total number of words and rate of response revealed similar differences between the melodic and spoken groups. Whereas there was a marginal effect for total number of words for both groups ($p = .057$), this effect was largely attributed to the pre- and post-intervention gains for the melodic group, which were greater than for the spoken group. With regard to rate of response, although it took both groups significantly less time to produce 100 utterances in the post-intervention language sample (p <.05), children in the melodic group produced the utterances in a significantly shorter period than children in the spoken group, requiring half as much time than they did in the pre-intervention language sample (correlation coefficient of .994; $p < .01$). Children in the melodic group also experimented more with the target phrases by modifying, extending or transforming them.

As for the mean length of utterance (MLU), a marginally significant effect was found (p =.060) which was almost entirely due to the post-intervention gains in the melodic group. As well, the significant correlation that was found between pre- and post-intervention scores for MLU indicate that incoming MLU had an effect on the magnitude of the gains made.

These findings provide evidence of MIT as an effective method for stimulating verbal speech in the way it mirrors early language development by exploiting the prosodic (melodic) characteristics of speech. Implications for future research were addressed and applications for using MIT with young children were also discussed.

RÉSUMÉ

Cette étude a examiné l'effet d'une adaptation de "Melodic Intonation Therapy" (MIT) sur le langage verbal communicatif de jeunes enfants trisomiques. Huit enfants étaient jumelés selon la longueur moyenne de leur énoncés et divisés en deux groupes, le groupe mélodique et le groupe parlé. Tous ont reçu la même intervention individuelle pendant douze semaines. Par contre, la présentation des phrases cibles étaient chantées pour un groupe et parlées pour l'autre. Les données étaient recueillies des échantillons de langage, avant et après l'intervention ainsi que des réponses verbales produites durant chacune des sessions d'intervention. Les resultats ont démontrés des gains plus grands pour le groupe mélodique pour le nombre total des mots, la longueur des énoncés et le temps de production, suggérant ainsi l'effet positif de MIT. L'auteur discute des implications pour la recherche future et des façons pratiques de mettre à exécution le MIT auprès des enfants.

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

Raw data, including language transcripts, fieldnotes and session data forms are available upon request

LIST OF FIGURES

FIGURES Page

LIST OF TABLES

ACKNOWLEDGEMENTS

The author gratefully acknowledges all those who helped in the realisation of this study - to the late Helga Feider for introducing me to CHILDES (a computer-assisted language analysis program), to Nancy Ship for her invaluable clinical suggestions during the initial stages of this study, to Ellen Moss and Mary Maguire for their guidance and interest, to Loriann Lorch and Sonya Guerriero for their assistance in coding and analysing the language transcripts, to Janusz Kaczoroirski for his help in the statistical analysis of my data and to Benoit Massicotte for his meticulous preparation of the musical examples. I would also like to thank the Humanities Research Grants Subcommittee for their financial assistance and Ayala Manolson for giving me permission to adapt a language assessment instrument from her book "It Takes Two to Talk". I am indebted to Riva Heft for preparing the graphs and for her immeasurable support throughout my years as a graduate student. A special thanks to my supervisor, Professor Eleanor Stubley, for her ongoing support and guidance. I am forever grateful to my family for their love and patience, especially to Mike, Sean and Josh. Finally, very special thanks to the children and parents, who participated in this project, for their cooperation, their enthusiasm and for making this study possible.

CHAPTER I

INTRODUCTION

Language is the source of human life and power (Chapey, 1994). The ability to communicate effectively is essential to every aspect of modern life and the key to social and economic independence (Nadel, 1992). From its origin in a baby's first cry, the acquisition of speech is a truly remarkable achievement. The child's developing language is based on the convergence of cognitive, sensorimotor and affective processes (Miller et al, 1980) and is influenced by the cultural and interactive forces in the social environment (Vygotsky, 1986; Conti-Ramsden & Snow, 1990; Bloom, 1993). If one or more of these developmental processes is not intact, there may be a delay in acquiring speech. Young children who are speech-delayed are likely to experience failure and considerable frustration at not being able to meet the expectations and hopes of others, especially parents. This can cause considerable frustration which might have serious consequences on their future emotional, cognitive and social development.

Children with Down syndrome are a case in point. Described by Langdon Down in 1866, Down syndrome is a major cause of mental retardation and congenital heart disease, occurring in approximately one in every 600 live births (Cicchetti & Beeghly, 1990; Kronenberg et al, 1992). Also known as Trisomy 21 because of extra genetic material on the twenty-first chromosome, Down syndrome is characterised by particular facial and other physical features, as well as by defects of the immune system (increased immunity to infection). Advances in the field of human molecular genetics are making it increasingly possible to discover the genetic basis for the associated defects, perhaps eventually preventing and treating them (Kronenberg et al, 1992).

Children with Down syndrome are also known to have a specific delay in sequencing words into speech patterns and diminished speech intelligibility (Fowler, 1990). Furthermore, as these children get older, their linguistic deficits may increasingly affect other areas of development (Cornwell & Birch, 1969). As Nadel (1992) notes, "the ability to communicate effectively is essential to social life, virtually all forms of gainful employment and just about every other aspect of modern life" (Miller, 1992, p.38). Speech language pathology treatment of children with Down syndrome has applied a developmental approach to improving speech intelligibility and speech-sequencing abilities. Certain techniques have reportedly been used, including sucking, chewing and swallowing to improve oral motor control, the use of signing and pictures to reinforce language comprehension and speech production, the practice of sound patterns, such as C-V-C (ex. phonemic drills) and the use of scripted events or structured child-adult interactions, including games, book-reading and role-play (Mahoney & Snow, 1983; Swift & Rosin, 1990; Spiker, 1990).

Although many areas of cognitive function in individuals with Down syndrome are of interest, the one area that has received the most attention is language. As a practising music therapist, what has often personally intrigued me during an initial assessment of children with speech delay is the ease with which they imitate complex rhythm patterns on a drum despite their inability to string more than two words together. These observations led me to explore further how the active ingredients of music, such as melody and rhythm, might improve speech sequencing abilities. There seemed to be a natural link between music and speech because of the elements they share - rhythm, melody, timbre, pitch, intensity, etc. When exaggerated within the context of musical activities, these elements facilitate vocal and verbal responses. More specifically, setting words to a melodic motif that reflected the intonation and rhythm of the speech pattern, then gradually increasing the length of the pattern, seemed to be an effective strategy in developing speech.

Melodic Intonation Therapy (MIT) was a method with which I was not entirely familiar, but one that was close to what I had already been using intuitively in my work (Carroll, 1989). It was first developed by Sparks, Helm and Albert in 1973 to aid speech recovery in adult aphasic patients at the Aphasia Research Center of the Boston's Veterans Administration Hospital. MIT is based on the Minor Hemisphere Mediation Model (Chapey,1994), that recognizes right (minor) cerebral dominance for music and speech prosody (Scheid & Eccles, 1975; Gates & Bradshaw, 1977; Goodglass & Calderon; Ross & Mesulam, 1979; O'Boyle & Sanford, 1988; Morton et al, 1990). By converting speech patterns to melodic motifs, MIT exploits the affective-prosodic qualities (or suprasegmental characteristics) of speech - pitch, loudness, rate and stress - to facilitate communicative speech.

Melodic intonation is a form of singing, dating back to the Judeo-Christian period (Sparks &

Deck, 1994) and is distinguished from speech by its slower, more lyrical tempo, more precise rhythm and more accented points of stress. Sparks and Holland (1976) noted that patients appeared to be more capable of processing the structural aspect of the intoned verbal speech patterns when they focused on the melodic line, rhythm and points of stress. This observation seems to be consistent with the current thinking of suprasegmental functions of intonation, rhythm and stress as the foundation or structural support for the organization of speech communication (Leung, 1985).

The original MIT protocol (1973) consists of four levels, gradually increasing in difficulty with regard to phrase length, and gradually reducing dependency on the clinician and reliance on intonation. At Level One, the process of intoning melodic patterns and handtapping the rhythm and stress of each pattern is established. At Level Two, the patient hums and taps the speech patterns together in unison with the clinician, then repeats the patterns after they have been modelled. Finally, the patient responds to a question, using the speech pattern. Level Three is similar to Level Two, except for an enforced delay of response[1], the purpose of which is to maximize efficiency of word retrieval. The aim of Level Four is to return to normal speech by way of the *sprechgesang* (speech-song) technique, in which the constant pitch of the intoned words is replaced by the variable pitch of speech, with the tempo, rhythm and stress of the speech pattern being retained. Throughout the procedure, patient and clinician sit facing each other. MIT gives the clinician flexibility in determining appropriate target phrases and in adapting to changes in intonation patterns created by the patient.

Sparks et al (1973) reported MIT's effectiveness in the recovery of communicative speech in three adult males who had lost the ability to speak, following left hemispheric damage. Six other studies have been found that provide evidence of the effectiveness of MIT. Two studies[2] examined the use of MIT in the speech rehabilitation of adult aphasics (Sparks et al, 1974; Marshall & Holtzapple, 1976). Four case studies reported the successful application of modified versions of MIT in the development of speech in children with language delays (Miller & Toca, 1979; Helfrich-

---

[1] The patient is required to wait a certain period before repeating the pattern.

[2] A recent Medline database search revealed an abstract of a third study, that reported MIT's effectiveness in the speech rehabilitation of 80 Romanian aphasics. Unfortunately the article was published in the Romanian Journal of Neurology and Psychiatry, and thus unavailable.

Miller, 1980; Romski, 1980, Krauss & Galloway, 1982). In addition, two articles provided a detailed description of MIT (Sparks & Holland, 1976; Sparks & Deck, 1994) and two papers examined criteria for candidacy for MIT (Berlin, 1976; Naeser & Helm-Estabrooks, 1985).

On the basis of its methodology, candidacy and efficacy, MIT has been rated as promising[3] by the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology (1994). As such, this melodic-based language stimulation method seems to be a suitable intervention strategy with a strong theoretical basis that meets the child's need for structure, appropriate communicative speech patterns and multisensory stimulation (Preuss & al, 1987). Furthermore, children with Down syndrome possess certain characteristics considered favourable with regard to candidacy for MIT (Sparks et al, 1974; Sparks & Deck, 1994), including difficulties in speech production as compared to language comprehension (Fowler,1990; Miller, 1992) and poor vocal imitation skills (Preuss et al, 1987).

Determining what kinds of intervention strategies are most effective in overcoming language difficulties is a top priority for researchers (Nadel, 1992). Whereas it was noted by the Assessment Subcommittee of the American Academy of Neurology (1994) that, as one of the few language therapy techniques formal enough to be evaluated, "MIT can fulfill consistency requirements for research-level studies" (p.566), there have been methodological flaws in the research done to date. The most apparent one has been the lack of a control group to account for time, maturation and practice (or carry-over effect). There have been no comparative studies done to test MIT's effectiveness. Furthermore, no studies have been found that examine the language acquisition of children with Down syndrome across different intervention strategies. Therefore, a study of the effectiveness of MIT in increasing the communicative speech of young children with Down syndrome would fill a void in the research literature.

---

3 Promising is defined here as "given current knowledge, this technology appears to be appropriate for the specified patient population" (referring to adult aphasics) (p.567, 1994).

STATEMENT OF PURPOSE

The purpose of the present study was to examine the potential effectiveness of Melodic Intonation Therapy (MIT) in improving the communicative speech of young children with Down syndrome. Chapter II will develop a theoretical basis for the study and will explore the issues regarding the implementation of an appropriate research design for understanding MIT's effectiveness with children with Down syndrome. Chapter III will describe the methodological procedures that were carried out and the rationale for using them. Techniques for data collection and analysis will also be explained. Chapter IV will present the results of the analyses of data and examine certain contextual factors that might have accounted for these findings. Finally, Chapter V will present a summary of the study, followed by a discussion of the findings, conclusions and implications for future research and clinical practice.

CHAPTER II

REVIEW OF THE LITERATURE

The aim of this chapter is to review the research literature that has provided both theoretical and methodological contexts for the present study. The literature will be reviewed here under the following three headings: Right hemispheric processing of speech and music, Music-based activities to develop speech, and Melodic Intonation Therapy.

RIGHT HEMISPHERIC PROCESSING OF SPEECH AND MUSIC

There is considerable evidence, primarily from brain lesion studies and dichotic listening experiments[4], to suggest that the right hemisphere may be dominant for the processing of speech and music prosody. Prosody, first described by Monrad-Krohn in 1947 as the "third element of speech" (along with grammar and vocabulary), is defined as "the melodic line produced by the variation of pitch, rhythm and stress of pronunciation that bestows certain semantic and emotional meaning to speech" (as reported in Ross & Medulam, 1979, p.146). Ross and Medulam (1979) described the case of two patients with right hemispheric lesions, who spoke with monotone voices that were devoid of the prosodic-affective qualities of speech, supporting the notion that prosody is a dominant function of the right hemisphere.

In a dichotic listening study by Goodglass and Calderon (1976), all subjects - 16 music students, demonstrated a left ear advantage for tonal stimuli, and a right ear advantage for verbal stimuli (words superimposed on piano notes). These findings led the authors to speculate that the phonological and semantic aspects of normal speech perception are processed in the left hemisphere, and the intonational contours in the right hemisphere. Similarly, O'Boyle and Sanford (1985) found a right hemispheric superiority for melody and a left hemispheric superiority for rhythm in a study of 46 male university students, who were asked to first listen to a familiar melody, then to determine whether the rhythmic sequence that was tapped in either their left or

---

[4] In these studies, the subject receives different auditory stimuli simultaneously, one to the right ear and the other one to the left ear.

right palm matched the rhythm of the previously presented melody. In a more recent study by Morton, Kerschner and Siegel (1990), sixteen males, aged 10 to 12 years, were asked to remember a series of digits under two different conditions, prior exposure to music or to silence. Prior exposure to music resulted in a significant increase in total digits reported. The authors concluded that music might be used to improve attention and memory, by increasing bilateral cerebral arousal levels, possibly through the mediating role of the right hemisphere. This finding is particularly important when considering the application of MIT to young children with Down syndrome in that the interest aroused by the melodic element may be a motivating factor in eliciting verbal output.

MUSIC-BASED ACTIVITIES TO DEVELOP SPEECH

Sound evidence of a neurophysiological link between music and speech, as illustrated above, has important therapeutic implications for using music to improve speech performance. Indeed, for more than four decades, music therapists and speech therapists have been investigating the effect of music on speech production. Studies have reported improved speech (i.e. increased vocal range, rate of speech and intelligibility) as a result of specific music activities, including the singing of familiar and specially composed songs, rhythmic chanting, telling stories set to music and the use of rhythmic and singing instruction (Seybold, 1971; Michel & May, 1974; Roger & Fleming, 1981; Leung, 1985; Hoskins, 1988; Cohen & Masse, 1993). Michel & May (1974) described three research projects (Marsh, 1969; Irwin, 1969 & 1971), which dealt specifically with speech production. In one of these projects (Irwin, 1971), visual cue cards, nonsense syllables and specially composed songs were used to help five nonverbal children with Down syndrome practise three speech sounds [b], [p] and [m]. Although the author reported an increase in the intelligibility of these sounds, the strength of these results is hard to determine. The children acted as their own controls, making it difficult to separate treatment effects from carryover or practice effects.

Another study done about the same time by Seybold (1971) sought to determine the effectiveness of a technique that combined music activities (particularly the singing of familiar songs) with conventional speech therapy techniques. Eight preschool "speech-delayed" males were treated individually for a period of eight weeks. Four children were exposed only to speech therapy techniques. The remaining four children were exposed to the combined music and speech therapy techniques. Due to a theft of the pretests, there was no baseline, thus making it impossible to evaluate the effects of treatment. There were other apparent flaws in this study. Firstly, the author did not clarify what was meant by "speech-delay". Similarly, the dependent variable of interest, spontaneous speech was not clearly defined. Therefore, it was not known exactly what kind of information the investigators intended to obtain. Thirdly, there was an incongruity between the dependent variable of interest and the dependent measure -using the Houston Test to measure spontaneous speech was questionable, given that this test is used to measure various aspects of language concepts, not spontaneous speech. Finally, the two treatments lacked standardized and theoretically sound protocol.

The study by Hoskins (1988) used a pre-posttest within- subjects, repeated-measures design to examine the effect of antiphonal singing[5] on the expressive language abilities of sixteen preschool children (mean IQ of 73). Four tests were administered: a pre-recorded musical imitation task and three versions of the Peabody Picture Vocabulary Test (PPVT, EOWPVT- Expressive One-Word Picture Vocabulary Test and Form M, the Melodic version). Following the pretests, children were divided into three groups on the basis of chronological age and functional abilities. After ten weeks of three weekly 30-minute sessions, a significant improvement was found only for the melodic version of the PPVT ($p<.05$), suggesting that antiphonal singing was beneficial. Despite reported gains, as in the Seybold study, there were serious problems with the research design. Firstly, the dependent variable of interest, expressive language abilities, was not clearly defined. Secondly, there was an incongruity between the dependent variable of interest and the test used to measure it. More specifically, the Peabody Picture Vocabulary Test (PPVT) did not seem to be a suitable tool to measure improvement in expressive language, particularly since the required response was either nonverbal (pointing to a picture, as in the PPVT) or a single word (Expressive One-Word PPVT). Thirdly, the lack of a control or comparison group made it difficult to determine the effectiveness of treatment. Finally, the probability of error was increased by the multiple statistical tests used to analyze the data.

A study by Rogers and Fleming (1981) involved the use of a music-speech therapy technique with a 53-year-old aphasic male. Where previous studies used subjects as their own controls in successive treatments, Rogers and Fleming followed their subject concurrently for four

5 Antiphonal singing refers here to a procedure in which the therapist showed a picture card to the group while singing a three to five-word phrase about it. The group then repeated the object name on the card with the therapist.

months in both music therapy and speech therapy. Findings revealed that every stage of speech recovery, including automatic speech for counting, the development of speech patterns and appropriate verbal response to questions, first appeared in the music therapy sessions. It was concluded that this increased rate of progress in music therapy might have been attributed to the music-speech therapy technique, which was characterised by the use of a single "carrier melody" ("Yankee Doodle") for all sentence items.

Whereas Rogers and Fleming (1981) intentionally used a familiar tune "to take advantage of any automatic ability present" (p.34), Sparks et al (1974) observed that the use of an intoned utterance resembling a familiar song often produced less than successful results, because the familiar melody often stimulated recall of the more intact non-communicative song lyrics. Subjects would experience difficulty focusing on the "unfamiliar" words, not traditionally associated with the tune. Based on these observations, they argued that there was a dramatic difference between the subjects' automatic speech of well-memorized songs and the deficient quality of their meaningful, communicative speech. Moreover, they challenged the presumptions of music therapists and speech therapists, that the singing of familiar songs could improve language skills in adult aphasic patients.

MELODIC INTONATION THERAPY (MIT) RESEARCH

<u>MIT research with adult aphasics</u>

Melodic Intonation Therapy (MIT) was first introduced in 1973 by Sparks, Albert and Helm. MIT involves setting communicative phrases, such as "May I have some juice, please" to melodic motifs that reflect the intonational contour and rate of stress of the speech pattern. As described in Chapter I, MIT has a well-defined protocol and is grounded in a neurophysiological theory that recognises right hemispheric dominance for music and speech prosody.

Two studies have provided evidence of MIT's effectiveness in recovering speech with severely speech-impaired right-handed adult aphasics. Sparks, Helms and Albert (1974) were the first to provide evidence that improvement in communicative speech occurred as a result of MIT. The authors used a two-treatment within-subjects design in which each patient acted as his own control. Six months of "traditional" language therapy, during which no progress was reported, was followed by an experimental period of daily individual MIT sessions and group sessions (a less structured form of MIT, where verbal interactions were intoned). Of the eight patients, six showed significant improvement in all subtests of the Boston Diagnostic Aphasia Examination (B.D.A.E.) (1972), with the most significant change occurring for phrase length. Four of these patients began MIT treatment with limited and meaningless but well-articulated stereotype jargon. By the end of treatment, they were using three or four-word phrases. The other two patients began treatment with a few overlearned social phrases and were using one or two-word phrases at the end of treatment. Finally, the two patients who did not improve were the most verbal, scoring very well in the pretests of repetition when compared with the other patients and with their own performances on the other verbal tests. It was concluded that MIT treatment was most appropriate for those with severely restricted verbal output, good auditory comprehension and poor verbal imitation skills. The authors also found that syntactic growth began to appear post-MIT, inferring that the benefits of MIT might be delayed. The strengths of this study lie in the use of an appropriate test instrument, the B.D.A.E., to examine different aspects of verbal expression (i.e. responsive naming, confrontation naming, phrase length) and in the establishment of criteria for MIT candidacy. An apparent weakness was the two-treatment within-subjects research design. In the absence of a control group, it was difficult to separate the effect of treatment from the carryover effect.

The three case studies by Marshall and Holtzapple (1976) provided further evidence of the delayed effect of MIT, but in this case it was with regard to speech intelligibility, rather than phrase length. Progress was noted on the basis of the Porch Index of Communicative Ability (PICA) (1974), with the most improvement occurring three months and six months post-MIT. In this study, the MIT protocol was simplified in the following ways to meet the needs of those patients, who were not responding to the complexities of the "orthodox" form of MIT: three treatment "plans"[6] were used instead of four levels, "intoned sequence units" served as the carrier phrase for several speech patterns (i.e. "It's a _______" and a core of five nouns) and graphic representations of the actual intoned target phrase were shown while practising the target. The authors clearly stated that their modified treatment was not meant to replace MIT, but only to

[6] Plan I comprised a series of steps, increasing in difficulty as in the original MIT protocol except for the return to normal speech via chanting; Plan II was similar to Plan I, but stressed the development of the patient's vocabulary; Plan III involved independent practice using a Language Master Machine.

respond to the specific needs of those patients who were referred to their clinic. This study is also noteworthy for its detailed descriptions of the progress of each patient at each level and of the verbal instructions that were provided. As well, the authors addressed the need for considering other factors influencing verbal output, in particular, the potentially critical role of handtapping in MIT treatment success.

MIT research with children with speech delay

To date, there have been only four case studies; Miller and Toca (1979), Romski (1980), Helfrich-Miller (1980) and Krauss and Galloway (1982). The issue of modifying MIT to meet the developmental needs of children for multisensory stimulation and for dynamic, pleasurable interactions is particularly significant in these studies. Miller and Toca (1979) described the case of a three-year-old nonverbal boy with autistic features who began producing spontaneous verbalizations as a result of MIT. The boy acted as his own control, having shown no progress after being treated for one year with a Simultaneous Communication method (signed and verbal language). The ten-week experimental period consisted of four weekly sessions. In each session, a cookie (or cracker) and a drink were placed on a tray in front of the boy. The researcher sang the target items three times with the appropriate signed gestures, followed by an intoned request "What do you want ?". After 25 sessions, the boy began to consistently sign and intone the three target items. By the end of treatment, he was spontaneously intoning other words at home, or on the school bus. His mother continued to use the MIT procedure at home and after 35 days, he was producing up to four-word utterances. Based on these findings, Miller and Toca recommended prolonging MIT treatment for at least three months. This recommendation echoed that of Sparks et al (1974), who advocated a minimum of three months of daily post-MIT. It must be noted that using food as a reinforcer was a confounding variable that seriously jeopardized the validity of these findings.

Romski (1980) provided an interesting way to meet the young child's need for multisensory stimulation by using a puppet to facilitate handtapping. Her study of a five-year-old apraxic child involved six months of a traditional treatment approach, during which minimal progress was made, followed by six months of MIT. Gains were reported with respect to the intelligibility of spontaneous two-word phrases and to their generalised use at school and at home. Improved speech intelligibility and phrase length were also reported by Helfrich-Miller (1980) in their study of two apraxic children in which American Sign Language was used to assist speech production.

In a study of two right-handed apraxic boys, Krauss and Galloway (1982) went even further to modify the MIT procedure in order to accommodate children's developmental needs. The first level of MIT was extended to give the children time to establish the intonation pattern. Signed gestures, puppets and pictorial representations were used to enhance the meaning of the target phrase and to cue the child's attention to the phrase. Each child served as its own control, with two months of traditional speech therapy followed by an experimental period where MIT was used as a warm-up (facilitating procedure) for 20% of the language therapy time. Gains in noun retrieval, phrase length and verbal imitation tasks were similar to those reported by Sparks et al (1974).

In addition to the MIT modifications, this study is also noteworthy for its operationalized definitions and clear rationale for the two test instruments used, Language Sampling and the Porch Index of Communicative Ability in Children (PICAC) (1974). Language Sampling, or the sampling of a child's language within the context of spontaneous interactions in naturally occurring situations, was used to measure each child's mean length of utterance (MLU). It has since become a major component of both clinical and research assessment procedures (Bloom & Lahey, 1978; Duchan & Lund, 1983) and has been compared favorably to standardized language tests (Blau et al, 1984). The use of the verbal and auditory sub-tests on the PICAC was justified on the basis of its similarity to the sub-tests of the Boston Diagnostic Aphasia Examination and to tests of repetition used by Sparks et al (1974). Finally, the use of a two-treatment within-subjects research design made it difficult to determine the effect of treatment. This was the case for all the MIT studies.

CONCLUSIONS

The few MIT studies that have been done to date all share two important points: 1) all the studies have been carried out with small groups or with a single subject, and have involved adult aphasics or children with speech-delays, and 2) all have been marred by methodological problems. With the exception of the study completed by Rogers and Fleming (1981), every MIT study used a two-treatment within-subjects design in which there was a period involving traditional language therapy, followed by an experimental period involving MIT. This made it difficult to determine whether reported gains were a result of treatment or practice. Sparks (1974) was very aware of the problems inherent in this research design, specifically with regard to carryover effects, such as time, maturation and practice. However, from an ethical point of view, depriving some patients for the purposes of research control was for him not easily justified or explained. Thurman and Widerstrom (1990) suggested that the issue of denying service might be resolved if researchers were to randomly assign children to different intervention programs and then collect data to examine which is more effective. If one is to effectively examine MIT, a comparison group is needed. This group would receive a treatment similar to the experimental group in all respects, except for the absence of the melodic component. In this manner, the potential contribution of this element in improving speech might be determined.

The issue of modifying MIT when working with children was also raised in the above literature review. If one is to effectively meet young children's developmental needs (i.e. social, emotional and cognitive), it would be important to allow for stimulating and dynamic interactions through careful choice of play materials.

Only one study examined the intervention process by noting the subjects' verbal responses at every level during each session (Marshall & Holtzapple, 1976). This raises the issue of the importance of examining the unfolding of the intervention period, including the contextual factors affecting change in verbal output, if one is to gain a better understanding of the MIT treatment process itself. Furthermore, collecting data from the intervention process and linking these findings to outcome measures might help to strengthen internal validity. It might also assist clinicians in developing effective strategies for the implementation of an adapted form of MIT with young children, and help researchers in refining methodological procedures.

RESTATEMENT OF PURPOSE AND SUB-PROBLEMS

The purpose of the present study was to examine the potential effectiveness of an adaptation of Melodic Intonation Therapy (MIT) in improving the communicative speech of young children with Down syndrome.

A number of questions were important in guiding the collection and analysis of data:

- How did the children's verbal output evolve during the 12-week intervention period?

- What differences in verbal output between the two groups could be identified:
  a) during the intervention process and
  b) at the end of the intervention period ?

- What factors affected verbal output during the intervention process and during the collection of language samples ?

- What recommendations can be drawn up for implementing MIT with young children ?

CHAPTER III

METHODOLOGY

SUBJECTS

Eight children (five boys and three girls) between the ages of three and six were selected to participate in this study. Three restrictions were imposed on subject selection: diagnosis of Down syndrome, production of at least one-word utterances, and English as principal language of communication at home (see Table 1).

Table 1
Subject descriptive data

	Child	Sex	Age yrs;mo	Incoming MLU	Diagnostic Category	School	Traits (personal & medical)
Melodic Group	A^M	F	5; 7	1.59	Down syndrome	Integrated kindergarten	social, playful
	J^M	M	3; 7	1.39	Down syndrome	Integrated pre-school	reserved, attentive; possible hearing loss
	L^M	M	6; 7	1.12	Down syndrome	Special school	friendly, resistive in a playful way; apraxic
	T^M	M	3; 10	1.26	Down syndrome	Integrated daycare	social, playful; possible hearing loss, many sinus and ear infections
Spoken Group	E^S	F	4; 1	1.37	Down syndrome	Integrated daycare	restless, curious
	J^S	M	3; 5	1.18	Down syndrome	Integrated daycare	social, gentle,caring
	R^S	M	3; 5	1.13	Down syndrome	Integrated pre-school	outgoing, active
	V^S	F	4; 9	1.17	Down syndrome	Special school	restless, imaginative; mild hearing loss (wears hearing aids)

Children were recruited with the help of several speech-language pathologists, a family doctor, a school nurse and parent-members of local associations for the handicapped, to whom letters were sent, explaining the study and requesting that they distribute copies to the parents of children with Down syndrome. Announcements were also placed in the newsletters of the Montreal Association for the Intellectually Handicapped, the West Island Association for the Handicapped and the local Down Syndrome Association. Prior to the study, parents were informed of the procedures and asked to sign a consent form which stated that their children could be withdrawn at any time during the study upon request (see Appendix 1 - Forms and Letters).

RESEARCH DESIGN

A pretest-posttest comparison group design was used. Three dependent measures - total number of words (also expressed as amount of verbal output), production time or rate of response (as determined by the number of minutes required by the child to produce 100 consecutive utterances) and mean length of utterance (MLU) - were obtained from pre- and post-intervention language samples collected during visits to the children's homes. MLU is a number representing the mean length of a continuous sample of utterances measured in morphemes[7] (Brown, 1973). There is evidence to support MLU as a useful summary measure of syntactical complexity (Cicchetti & Beeghly, 1990) and as a reliable measure of a young child's language competence, when compared to standardised tests, such as the Carrow Elicited Language Inventory (CELI) (1974).

The children were matched on the basis of incoming MLU scores and randomly assigned to one of two groups. Although not used as a basis for matching groups, subject factors, such as physical and emotional state at time of testing, hearing status, specific language difficulties, sociocultural influences, cognitive function and chronological age, were considered to facilitate the development of specific intervention techniques and for data analysis and interpretation.

INTERVENTION

---

[7] A morpheme is the minimum unit of meaningful speech. It can stand alone (free morpheme), such as the word "cat", or it can be attached to a word (bound morpheme), such as "s" in "cats"; the "s" provides information regarding quantity.

The intervention period consisted of 12-weekly 30-minute individual sessions that were carried out by myself within the context of a core set of interactive play situations. Each child in the melodic group (N=4) received a modified version of Melodic Intonation Therapy (MIT). Each child in the spoken group (N=4) received the same treatment as the melodic group in all respects, but without the melodic component. More specifically, the only difference between the two intervention strategies was the manner in which the models and requests were presented - all the models and requests were melodically intoned for children in the melodic group, whereas they were spoken for children in the spoken group.

The two intervention strategies were modelled on the basis of the original MIT protocol (1973) and were adapted in the following ways to meet the developmental needs of young children with Down syndrome for stimulating, active and playful interactions:

1. Three levels of response were used instead of four. Each of the three levels of the intervention corresponded to a certain type of response. At Level One, after eliciting the target word or phrase, using a variety of linguistic and nonlinguistic cues, I invited the child to intone or say the target with me (unison response). At Level Two, I asked the child to repeat the target after it was modelled (imitative response). Finally, at Level Three, I asked a question to which the child was expected to respond appropriately with the target or other phrase (conversational response). The order in which these responses were elicited was variable, determined by the way in which each session unfolded.

2. The return to normal speech by way of *sprechgesang* (speech-song) and chanting was not considered, as it would have applied only for the melodic group. Ensuring that both groups received the same treatment throughout, except for the manner in which the models and requests were presented, was the top priority.

3. Whereas in the original MIT protocol, to progress through the various steps of each level, the patient's score had to be 90% based on the mean of 10 consecutive scores, the children in the present study were simply required to give three correct responses (or approximations) before moving to the next target phrase or level of response. The scoring

procedure was simplified to account for the relatively short attention span of young children - requiring 10 consecutive responses before moving on to the next target or level might have provoked resistance and uncooperative behaviour.

4. Body actions, pictures and actual objects were used to enhance the meaning of the target word or phrase, and to stimulate children in different ways - auditory, visual, tactile and kinaesthetic. For example, the target phrase, "I see dog" was practised while pointing to the picture of a dog, that was in a book or on a separate piece of paper. Similarly, phrases such as "stand up" and "sit down" were accompanied with body movements. The physical rigidity of patient and clinician sitting at a table facing one another would have been inappropriate for young children, given their need for movement and varied stimulation.

5. A bongo drum was used, instead of handtapping, to emphasize the rhythmic pattern of the target phrase as well as to facilitate listening and turntaking. Occasionally, body parts were used rhythmically instead of the drum (i.e. tapping index finger on target object, gently nodding foreheads together).

6. Hand puppets ("Super Bunny" and "Peter Parrot") were used to encourage role-playing. For example, "Super Bunny" was used by the child or myself to model, request and elicit unison responses, often while playing the drum. "Peter Parrot" was used primarily at the end of the session to praise the children for their efforts during the session and to say good-bye (see Appendix 2 for a description of the adapted MIT protocol used in this study).

IMPLEMENTATION OF INTERVENTION

DETERMINATION OF TARGET PHRASES

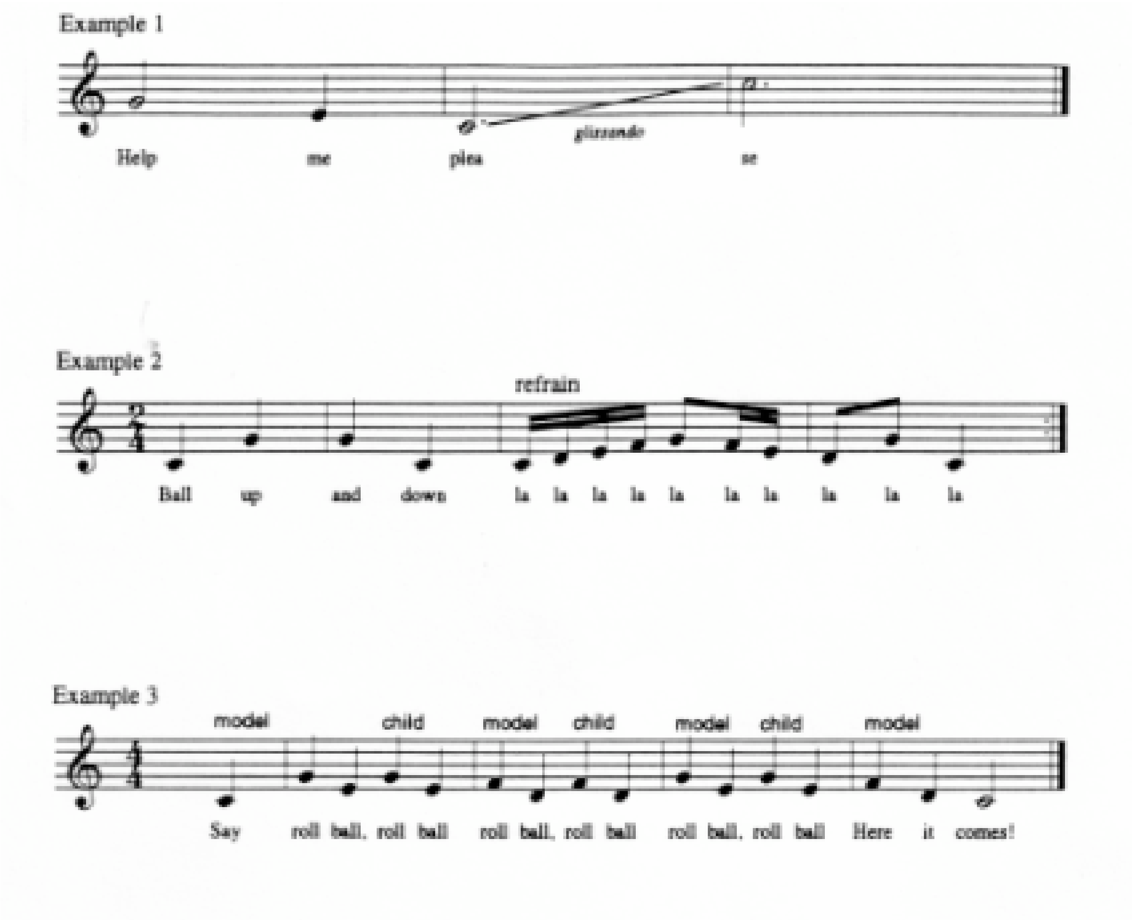
Target words and phrases were chosen on the basis of parent reports (Appendix 1), pre-intervention language samples and by the play materials used in the session. In addition, single-word and multi-word utterances were selected from among those commonly used, according to child language acquisition data. For example, single-word utterances include <u>objects</u>, such as names of food, toys and body parts, <u>actions</u> - "go", "sleep", "open", "eat", <u>social words</u> - "bye-bye", "yes", "no", and location words - "there", "up", "down". Multi-word utterances include action + object - "eat apple", "throw ball", agent + action + object - "mommy eat apple", action + object + location - "throw ball up" (Bloom, 1991).

MELODIC INTONATION OF TARGET PHRASES

As stated above, the only element that distinguished the treatment received by the melodic group from that received by the spoken group was in the manner in which target words and phrases were presented. In the melodic group, models and requests were melodically intoned. Target phrases were set to melodic patterns that exaggerated the prosodic elements of the speech pattern, including the intonational contour, rhythm and rate of stress. The melodic range was generally small, however, sometimes I would gradually rise in pitch to the octave (i.e. C→ C^1) to emphasize the last word of a phrase (Figure 1, Example 1). Occasionally, I would add a refrain as a link to a repeat of the target phrase (Figure 1, Example 2) or create a song that included imitative or echo responses (Figure 1, Example 3). Familiar motifs were avoided.

Figure 1. Selected melodically intoned

phrasess



ELICITATION TECHNIQUES

Particular target words and phrases were elicited within specific play contexts, gradually increasing in length. The sequence "drum" → "play drum" → "I play drum" was used while playing the drum; "help" → "help please" → "Debbie help please" was used when the child needed help (i.e. to remove the lid of the plastic container with the playmobile figures). Figure 2 illustrates how certain target words and phrases were elicited while playing with a ball, first with children in the spoken group then with those in the melodic group.

Figure 2. Techniques used to elicit target word or phrase

SPOKEN GROUP

NL I bring out ball
L.1 "What is this?"
L.2 "This is a (b*)".
L.3 Say "ball"

NL Get ready to roll ball
L.1 "What do you want me to do?"
L.2 "I'm going to (r*)".
L.3 Say "roll ball"

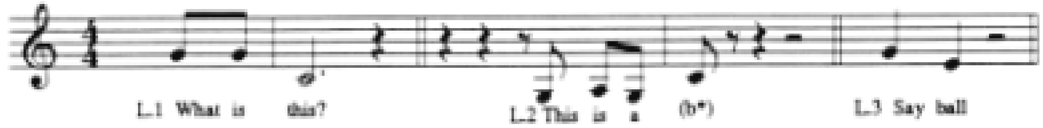
NL Put ball in basket. Shrug shoulders and show palm of hands.
L.1 "Where is the ball?"
L.2 "The ball's ____ "
L.3 Say "Ball allgone " or "No more ball"

NL = Nonlinguistic Cue; L.l = First linguistic cue, etc.

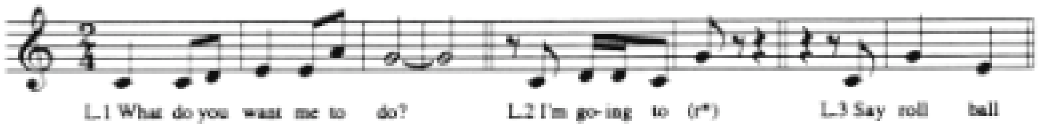
*Put lips together to form the consonant sound

MELODIC GROUP

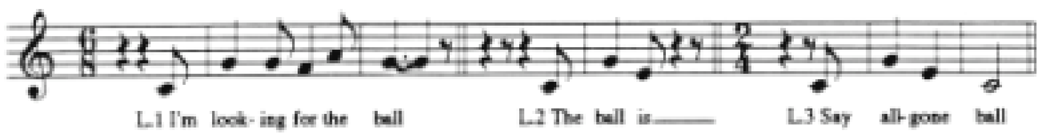
NL I bring out ball



NL Get ready to roll ball



NL Put ball in basket. Shrug shoulders and show palm of hands.



PLAY MATERIALS

Play materials included a ball (i.e. rolling, throwing, and hiding) for stimulating physical movement. Playmobile figures (i.e. taking a walk, eating a snack, going to sleep and waking up) were used for stimulating imaginary play with the number and type of figures usually corresponding to the people in the child's immediate family. Several children's books and a series of pictures (of animals, food and play materials) were used for identifying objects. A bongo drum (with 5" and 3" drumheads) was used to rhythmically support the target words and phrases, and two hand puppets, "Super Bunny" and "Peter Parrot", were used for encouraging role-play.

DESCRIPTION OF A TYPICAL SESSION

The unfolding of each session, including the choice and nature of the activity, was determined largely by the child. A session typically began with a greeting ("hi", hello Debbie", "How are you?", "I'm fine"). Play with figures might follow, which involved taking them for a walk and then giving them a snack. Next, the child might choose to play with the ball, throwing it up or hiding it in different places. The child might then take a book (or animal pictures) out of my toy bag, identifying the illustrations with the help of the drum. The session would typically end with "Peter Parrot", who praised the children for their efforts during the session and said "good-bye".

EQUIPMENT

A small Panasonic tape recorder (Model No. RQ-356A) was used to record each session. I chose to use a tape recorder because it was unobtrusive and easily hidden from sight, thus maximising the naturalness of the interaction.

ROOM

It was arranged that sessions would take place at the same time in the same room in order to control as much as possible for intra-session history. All room changes were unavoidable, although most were known beforehand (i.e. daycare was closed so child seen at home, or room was being renovated so session took place in child's bedroom). In two instances, the room change was unexpected, both for myself and for the child.

PILOT TEST

A pilot test, involving two normal 24-month old children (MLU between 1 and 2), allowed me to develop, refine and become familiar with the intervention strategies. This included determining the target phrases and associated intoned patterns, as well as practising the modelling and elicitation of spoken and intoned target phrases at the different levels of intervention. On the basis of the pilot test, it was decided that each session would be 30 minutes in order to allow for time to get to the room and to settle down. Certain observations made during the pilot test were quite different from prior expectations. For example, the move through levels 1 and 2 was faster than I had anticipated. As well, if I modelled a target word or phrase more than four times without getting a response, the children's attention and interest decreased. The introduction of the drum seemed to elicit quicker and more articulate responses, regardless of whether the phrase was spoken or intoned. This observation prompted me to introduce the drum earlier than I had originally intended, for motivational reasons. Finally, as a result of the negative response to “Peter Parrot" by one of the children, I decided to use him only at the end of the session for verbally reinforcing the children for their efforts and for saying good-bye.

PROCEDURES FOR DATA COLLECTION - HOME VISITS

COLLECTION OF LANGUAGE SAMPLES

The purpose of the home visits was to collect pre- and post-intervention language samples of each child at play in order to obtain the three dependent measures - total number of words, MLU and production time. Language Sampling was chosen as it is a major component of both clinical and research assessment procedures (Bloom & Lahey, 1978; Duchan & Lund, 1983) and has been compared favorably to standardized language tests (Blau et al, 1984).

Upon arrival at the child's home, I asked one of the parents to play with his/her child for fifteen minutes, after which I played with the child for the same period of time. This procedure was borrowed from Lund and Duchan (1988), who suggested that a preferred way to obtain a language sample was for the clinician to observe the child in interaction with the caretaker for 10-15 minutes, then to join in. All child-adult interactions were taped; however, the sample that was transcribed, and from which pre- and post-intervention measures were computed, contained only 100 consecutive child utterances. This number represented the minimum amount of child utterances needed to measure MLU (Brown, 1973). In order to ensure consistency and uniformity in the data collection procedures, each transcript included the last 30 utterances, produced by the child while playing with the parent (or sibling), followed by the first 70 utterances produced while playing with me (who presented the same play objects to each child - ball, figures, books, etc.). I chose to use the last 30 utterances produced by the child during the parent-child interaction to allow for a warming-up period; I chose to use the first 70 utterances while playing with me to get to know the children and to have a basis for comparing their language performance in interaction with me during the home visits with their verbal output during the intervention sessions.

Fieldnotes were written after each home visit while listening to the tape of the session. Particular attention was paid to the context in which the language sampling took place and its possible impact on the child's verbal output. Subject factors were also considered, such as the child's physical and emotional state at time of testing, hearing status, specific language difficulties and sociocultural influences.

TRANSCRIPTION AND CODING OF LANGUAGE SAMPLES

A computer package entitled Child Language Data Exchange System (CHILDES) (MacWhinney & Snow, 1984) provided the data input, storage and scoring capabilities for the language samples. CHILDES allows for greater scientific rigour in the collection, transcription and coding of data than doing it manually. Computerized transcriptions of language samples have enabled researchers to share data thus advancing child language research. MacWhinney (1991) estimates that there are over sixty groups of researchers all around the world who are collecting and transcribing language data using the CHAT system.

Language samples were transcribed by an independent observer with experience in using CHILDES and myself, according to CHAT format specifications (see Appendix 4). CHAT is the anacronym for Codes for the Human Analysis of Transcripts. Consensus (inter-rater) reliability on the content of the language transcripts was established by having the independent observer check 10%-15% of two of the pre-intervention transcripts that were prepared by myself. Percentage agreement was 85%.

PROCEDURES FOR DATA COLLECTION - INTERVENTION PERIOD

Data was derived from fieldnotes detailing the unfolding of each session, and from session data sheets documenting every response of the child.

Fieldnotes were written after each session while listening to the audiotape of the session. The focus of interest was the children's verbal output. In particular, the effectiveness of certain intervention strategies was noted as well as other factors affecting output. Among the factors identified were physical setting, child-researcher relationship, play routine, role of the drum, and manner of presentation of the target phrases (spoken or melodic).

In addition, an overview of each session was achieved by creating session data forms (Appendix 3). Following a review of the audiotape of the session, every verbal response or utterance of the child, as well as the play context in which it was produced, was noted on these sheets. Three categories were created to correspond with each of the levels of the intervention: unison response (target is intoned or spoken in unison), imitative response (target is repeated after researcher) and conversational response (target or other intoned/spoken phrase is produced in response to researcher's question. A fourth category, spontaneous response, was added in order to account for the utterances that were initiated spontaneously by the child.

While not corresponding exactly with the dependent measures for total number of words and mean length of utterance, the number of responses, collected during the sessions, provided an indication of the amount of verbal output, and therefore a means of comparing children's output during the intervention process with their output before and after it.

PROCEDURES FOR DATA ANALYSIS - HOME VISITS

SCORING OF DEPENDENT MEASURES

CHILDES has become a standard tool for accurate and reliable analysis (1984). Reliability for the scores for MLU and total number of words was ensured through the use of the CHECK program, which was run on each transcript file several times until no error messages were reported. The CHECK program verified data accuracy and the correct use of CHAT codes in preparation for the automatic scoring of the two dependent measures by the CLAN programs (CLAN is the anacronym for Computerized Language Analysis). The third dependent measure, production time, was calculated by listening to the taped language samples, each containing 100 consecutive child utterances, and determining the number of minutes it took to produce them in interaction with a parent or myself.

ANALYSIS OF DEPENDENT MEASURES

Two statistical tests were applied to analyze the three dependent variables - total number of words, mean length of utterance (MLU) and production time. A multivariate repeated measures analysis of variance (MANOVA) was calculated on each of the measures to determine whether the pretest-posttest gains between groups differed significantly. The Pearson product-moment correlation coefficient measured the degree of association between pre- and posttest scores.

PROCEDURES FOR DATA ANALYSIS - INTERVENTION PERIOD

ANALYSIS OF SESSION DATA FORMS

The number of responses or utterances within each category and the total number of responses for each session were computed. In addition, the relationship between the amount of verbal output at the different levels of intervention was examined. Increases in the length of the responses were also noted.

ANALYSIS OF FIELDNOTES

Fieldnotes were examined and coded for emerging themes in terms of the contextual factors that affected the children's verbal output. Among the factors identified were physical setting, child-researcher relationship, play routine, role of the drum, and melodic versus spoken presentation of the target phrases. When examining the setting in which the intervention took place, the physical (home, school, or daycare; spacious vs. cluttered), functional and auditory features of each setting were considered. The child-researcher relationship was determined by the evolving relationship and its impact on the child's verbal output. With regard to the play routine, elements such as the nature and preference of the play routine as well as the role assumed by the child (initiator, follower, partner) were of interest. The effect of the drum on the quality (i.e. clarity and rate of response) and quantity of verbal output was also considered.

Descriptive and quantifying data from the intervention period was generated in order to provide some answers to the following research questions: How did the children's verbal output evolve during the 12-week intervention period ? How did the children respond to the different levels of the intervention protocol ? What differences in verbal output could be observed between the four levels of response ? Was there a particular pattern of response that emerged ? What differences between the two groups could be identified ? To what extent did certain contextual factors, particularly the manner in which the target phrases were presented (spoken versus melodically intoned) have an effect on the quantity and quality of verbal output? What, if any, effect did the the drum have on the quality and quantity of verbal output ?

CHAPTER IV

FINDINGS

This chapter is divided into four sections. The first section addresses the factors affecting the collection of pre- and post-intervention language samples during the home visits. The second section examines the verbal output in the intervention sessions for both subject groups. The third section addresses the contextual factors influencing verbal output during the intervention process. The final section presents the pre- and post-intervention differences in verbal output, as it relates to the total number of words, mean length of utterance and production time.

In order to preserve anonymity, only the child's first initial will be used when making specific references: A^M, J^M, L^M and T^M are in the melodic group; E^S, J^S, R^S and V^S are in the spoken group. The superscript identifies the group to which the child belongs.

HOME VISITS

COLLECTION OF PRE-INTERVENTION LANGUAGE SAMPLES -

First home visit

EXCERPT 1: L^M

> I enter the small, living room with six- year-old L^M and his mother, followed by his father and older brother. A couch, two sofa chairs and a television are all the furniture that is in the room. The father picks up the lone ball lying on the floor, rolls it to L^M and says, "Give me the ball". L^M replies, "No", "no way". After several minutes, his brother is asked to get some toys. He returns with crayons and a colouring book with all its pages coloured in. The father asks L^M to "choose the green one" and so forth. L^M takes the appropriate crayon and, after scribbling some lines on the back of the colouring book, puts the crayon back in its proper place. Except for

several instances of "no" and "no way", L^M remains silent during this activity. His father often reminds him to "sit up straight" or "be good".

A few moments later, I sit on the floor next to L^M. His eyes are fixed on the bag of play materials that are at my side but he does not reach out for it. I show him "Peter Parrot" (hand puppet). L^M motions for me to put Peter on my hand. Resisting my attempts to engage him in play, he repeats "No" several times with a twinkle in his blue eyes. I put Peter away and take out a red nylon case containing a tambourine. L^M reluctantly unzips the case, shakes the tambourine for three seconds and, after refusing to play a vocal imitation game with me, says "dow(n)" (he means "put away"). When I ask him to put the tambourine in the bag, he refuses and continues to shake it. "Put it away, c'mon", his father echoes my request.

Ten minutes later...

We look at a book together. L^M is interested. Despite his apparent difficulty in repeating certain sounds due to his apraxia (see glossary), he says "book", "apple", "cocoa", "cup", etc.

EXCERPT 2: A^M

Colourful fabric highlights the narrow hallway that leads to the living room. I follow 5 year-old A^M, her sister (ALY), and her sister's friend into the small, living room, filled with pictures and tapestries. Birthday cards and small figurines adorn a shelf. A^M's mother, father and 16-year old brother are already sitting on one of the two sofas. The girls move the coffee table so they can sit on the floor. A^M's sister is holding a pack of cards."Play cards okay?", A^M suggests. Her sister asks, "Which game do you want to play ?" "War", A^M quickly replies. There is a lot of laughing and playful teasing. Within this context, A^M produces a steady flow of phrases, especially commands, such as "do this", "ALY come here". When the game ends, I take the ball out of my bag and join in the play. "Do you know what this is ?", I ask. A^M responds, "Play ALY.." We all play together. A^M repeatedly tells us what to do.

Analysis:

These two excerpts, taken from my fieldnotes during the first home visit, show striking differences, particularly in the way the father and sister interacted with L^M and A^M respectively. L^M's father was directive in his approach and concerned with the proper way of doing things. He chose the play contexts (i.e. play with ball, colouring) and did not give L^M an opportunity to assume the role of initiator. Nor did he praise L^M or offer words of encouragement. In contrast,

A^M's sister interacted with A^M on equal terms in a non-directive manner.

These two interactional styles were encountered in the home visits for both groups. Where the directive approach was observed with other children, the negative impact was not as pronounced as it was with L^M. The most common interactional style was one that combined a non-directive and directive approach (see Table 2).

Table 2
<u>Interactional style of parent during collection of pre-intervention language sample</u>

	Location	Parent (or sibling) at play with child	% (approx) of directed exchange	% (approx) of non-directed exchange
L^M	living room	father	90	10
J^M	den	mother	50	50
T^M	den	mother	50	50
A^M	living room	sister	25	75
E^S	daycare	educator	50	50
J^S	playroom	mother	50	50
V^S	living room	father	60	40

It was interesting to note that, for those children exposed to a balanced approach (i.e. relationship between the amount of directed vs. non-directed exchange approximately 50-50), the language sampling took place in a playroom with lots of toys, each child interacting with their mothers, who used effective verbal elicitation techniques to stimulate play (i.e. J^S and his mother created a person out of "playdough", each taking turns adding one part until the body was formed). For those children, where the approach was primarily non-directive, the language sampling took place in a living room or den in the presence of other family members. These conditions existed for those that were exposed to a primarily directive approach as well. The one exception was E^S, whose language samples were obtained at daycare. This location was chosen because Hungarian was the main language spoken at home and her mother felt that speaking English at home might be confusing. She added that the daycare was E^S's "second home".

Differences in interactional style and play context had an effect on verbal output, particularly in the way it set up certain modes of response in the child. As previously seen in the opening excerpts, the directed approach of L^M's father limited L^M's opportunities to use different kinds of speech or to engage in a conversation. Instead he protested most of the time, often saying "no" and "no way" while colouring (non-interactional activity chosen and directed by his father). In comparison, A^M's bossiness while playing "war" with her sister elicited 2 and 3-word commands such as "play ALY", "pass it", "ALY, come here", "play cards, okay?".

These modes of response were present in both groups. In the case of J^S, collaborating with his mother to make a person out of "playdough" elicited phrases such as "my turn" and "help me please". For V^S, pretending she was a waitress (serving all kinds of plastic food and drink - doughnuts, chicken, hot dogs, coke, tea, etc.- to her mother and father) resulted in a variety of one-word utterances, such as "(ba)nana", "chi(ck)en" and "co(ke)". J^M pushed a plastic dinosaur back and forth to his mother. He was animated and clearly having fun. This interactive play resulted in a variety of spontaneous utterances, including "ready", "go", "(hoo)ray" and "move back". As for R^S, he was basically on his own, moving from one object to the next and speech-babbling to himself. Intelligible output was limited to "hi, jeejee" ("talking" on phone to his cousin). Sporadic attempts by his brother to direct R^S's attention were unsuccessful.

I noted in my fieldnotes that, not only did the interactional approach of the parent or sibling set up certain modes of response in the children, but these modes of response were carried over in the way they interacted with me. A^M played cards with me and told me what to do. V^S took my order and served me tea and doughnuts. L^M resisted my attempts to play with him and continued to produce his stereotypic utterances, "no" and "no way". It was only during the last part of the language sampling, while looking at a book together, that L^M began to move out of an uncooperative mode and produce for the first time a variety of words. This was the case for the other children as well, except for J^M and R^S, who altered their mode of response when I joined in the play. Whereas J^M was animated with his mother, he did not respond verbally to me for the first five minutes. He seemed withdrawn and distrustful. According to his mother, he did not feel comfortable with females because of numerous hospital experiences and unpleasant encounters with nurses.

Consequently, she did not feel that the language sample was representative of J^{M}'s present level of speech. As for R^{S}, I noted in my fieldnotes that his brother did not attempt to structure an interaction around R^{S}'s play preferences, nor make an effort to elicit any verbal responses. When I joined in, the change in mode of response resulted in the production of one and two-word utterances (i.e. "bag", "boy", "madee der"(= mommy there), "fall down"):

> The television is on but the volume is turned down. R^{S} rides his tricycle. His brother observes. After a short while, R^{S} rides to the plastic telephone and "talks" to his favourite cousin. He then gets off the bike, runs to a corner of the room, picks up a book and begins to "read". Less than a minute later, he is at the VCR, trying to insert a tape (Barney). While "talking" and "reading", his speech is babble-like and unintelligible. When I sit down on the floor, he quickly sits down next to me and looks in my bag of toys. He takes out the playmobile figures, playing with them for more than five minutes and responding verbally with 1 and 2-word intelligible utterances, including "daddy", "walk" and "sit down".

COLLECTION OF POST-INTERVENTION LANGUAGE SAMPLES -

Second home visit

Post-intervention language samples were collected under similar circumstances with contextual factors, such as physical setting, interactional style and play activities, having a similar impact on verbal output.

There were, however, some differences, as seen in Table 3 that seemed to affect verbal output. In the case of E^{S}, the change was in the person who played with her before I joined in. This educator's repeated question "What's this?" elicited only one-word utterances (i.e. "egg", "juice", "{j}ello") while labelling pictures on puzzle pieces. In comparison, the playful interaction that was established with the first educator during the collection of the pre-intervention language sample, elicited steady output and the occasional two-word utterance ("help please", "did it!").
E's limited output was also in contrast to the two and three-word utterances she produced with me (i.e. "help please, "dad sleep", "sit down mom") while playing with playmobile figures during the

post-intervention language sampling.

A difference between the first and second home visits was also seen with J^S and T^M. In the second visit, J^S's mother was ill. In my fieldnotes, I noted that J^S was unconcentrated and that his Mom had to prompt him several times before identifying flash cards. According to her, J^S normally did not require any cueing. When I joined in, he immediately took everything out of my bag, refusing to put anything back. This conduct was not typical of his behaviour during the pre-intervention language sampling or during the intervention sessions. During this visit, three or four minutes passed during which there was little or no verbal output in comparison to the first home visit, where there were no such gaps. As for T^M, he was in fine form and very responsive during the second home visit in contrast to the first one, where he had a bad sinus infection. His illness might have affected pre-intervention scores, possibly accounting for the fact that, next to A^M, he had the largest post-intervention gains compared to the other children.

Table 3
Interactional style of parent during collection of post-intervention language sample

	Location	Parent (or sibling) at play with child	% (approx) of directed exchange	% (approx) of non-directed exchange
L^M	living room	mother*	75*	25*
J^M	Den	mother	50	50
T^M	his bedroom	mother	50	50
A^M	parent's bedroom	sister	35*	65*
E^S	daycare classroom	different educator*	75	25
J^S	playroom	mother	50	50
V^S	living room	father	40*	60*
R^S	Den	brother	25*	75*

*This indicates that there was a change from the collection of the pre-intervention language sample.

The most important difference was that all the children participated in twelve individual weekly language stimulation sessions since the first home visit, resulting in increased familiarity with me, as well as with certain play routines, established during the intervention. The effect of the familiarity was seen particularly with regard to the ease with which the children moved from playing with their parent (or sibling) to playing with me. The one exception was L^M, who did not want to stop playing with his mother. He reacted negatively to my joining in ("no", "away you"). However, once his mother's continued presence was assured - it took approximately three minutes to decide where she should sit - L^M was able to settle down and play with me.

As for the effect familiarity with certain play routines had, it was interesting to note that two children, T^M and J^S, were quite upset when they realised that the drum was not in my bag, as was reflected by their sighing and saying "drum" several times. That response might be understood in light of the fact that the drum had become an integral part of the intervention process, particularly in the way it structured the play activity and the speech patterns, regardless of group. I had decided not to bring the drum to the second home visit so that procedures for data collection were consistent with the first home visit, where the drum was not used.

INTERVENTION PROCESS

In this section, changes in the children's verbal output seen during the intervention process with regard to total number of responses, length of utterance, levels of response and clarity of production will be presented. These findings will then be analysed within the social-interactive play context in which the intervention process unfolded.

AMOUNT OF VERBAL OUTPUT DURING INTERVENTION PROCESS

The total number of verbal responses per session increased for all but one child (V^S), regardless of group, as illustrated in Table 4.

Table 4
<u>Total number of responses per session</u>

Melodic Group

Session	1	2	3	4	5	6	7	8	9	10	11	12
A^M	84	73	84	66	75	78	115	149	__	118	165	88

J^M	42	75	70	52	93	39	43	79	104	118	159	123
L^M	87	137	160	184	145	81	126	221	152	__	108	171
T^M	122	52	104	112	90	90	98	111	126	__	131	133

Spoken Group

Session	1	2	3	4	5	6	7	8	9	10	11	12
E^S	102	117	107	78	158	107	95	142	120	167	130	149
J^S	__	57	89	98	122	128	126	132	152	160	152	158
R^S	96	77	86	72	63	64	71	111	73	136	180	58
V^S	145	134	111	128	153	161	162	134	174	102	88	117

N.B. The dashes indicate that the total responses were not available due to a problem with the tape recorder.

Of the four children who showed a steady increase, three were in the melodic group - A^M (session #4-#11), J^M (session #6-#12) and T^M (session #6-#12). The fourth child, J^S, was in the spoken group (session #2-#12). For three of the remaining four children, L^M, E^S and R^S, verbal output was more variable, although a general increase was apparent across sessions - L^M (87 responses in session #1 compared to 171 in session #12), E^S (102 responses in session #1 compared to 149 in session #12) and R^S (96 responses in session #1 compared to 180 in session #11). In the last session, R^S produced only 58 responses; an unexpected room change affected verbal output in that he did not respond verbally for the first ten minutes of the session. For the fourth child, V^S, verbal output was less in session #12 than in session #1. A slight but erratic increase until session #9 was followed by a significant drop in sessions #10 and #11. Session #11 was unusually short (20 minutes instead of 30), thus accounting for the low verbal output (88 responses).

To summarize, three of the four children in the melodic group showed a steady increase in output, whereas three of the four children in the spoken group showed an increase that was more variable. This resulted in a slightly greater average increase in verbal output for the melodic group than for the spoken group (84-129 responses for the melodic group compared to 96-120 responses for the spoken group).

LENGTH OF UTTERANCE

While there was a general increase in total verbal output, there was also an increase in the length of utterances in both groups, as shown in Table 5. Most children produced 1-2-word utterances in the first session. During the course of the treatment, of the three children who improved the most, two were in the melodic group, A^M and T^M. By the last session, A^M produced an average of 3-4 word utterances with up to 7-word phrases spontaneously; T^M produced an average of 2-word utterances and up to 4-word utterances. The third child, J^S, was in the spoken group. He produced an average of 3-4 word phrases in conversational responses and up to 7 words in imitative responses (ex. "I see a little boy brushing teeth", session #8) by session #12.

The child who improved the least was L^M, whose neurologically-based apraxia made it difficult to sequence words or even parts of words (syllables). Some improvement was noted, however. At first, he had much difficulty sequencing the two syllables in a word, such as "pee - tah" (peter) or "e - ee" (debbie). Throughout the course of treatment, he progressed from combining syllables to producing several two-word phrases by session #8 and even a few three-word phrases by session #12 (i.e. "no mama bye").

Table 5
Range of length of utterance (in words) during intervention process

Session	1	2	3	4	5	6	7	8	9	10	11	12
A^M	1-4	1-5	1-5	1-4	1-5	2-3	1-5	1-6	1-5	2-5	1-7	2-5
J^M	1	1	1-2	1-2	1-2	1	1	1-2	1-3	1-3	1-3	1-3
L^M	1	1	1	1	1	1-2	1	1-2	1-3	1-2	1-2	1-2
T^M	1-2	1	1-2	1-2	1-2	1-2	1-2	1-2	1-3	1-3	1-3	1-4
E^S	1-2	1-2	1-2	1-2	1-2	1-2	1-2	1-2	1-2	1-2	1-3	1-3
J^S	1-2	1-3	1-3	1-4	1-4	1-4	1-4	1-6	2-4	2-5	2-4	2-5
R^S	1	1-2	1-2	1-2	1-2	1-2	1-2	1-2	1-3	1-3	1-3	1-3
V^S	1	1-2	1	1-2	1-2	1-2	1-2	1-2	1-2	1-3	1-2	1-2

LEVEL OF VERBAL RESPONSES

An analysis of the session data forms revealed that two types of changes emerged with respect to the levels of verbal responses - unison, imitative, conversational and spontaneous (see Figures 3 & 4). First, regardless of group, the number of unison responses increased in the first few sessions then gradually decreased, never exceeding 20% of total responses during the course of the intervention. There were two exceptions. L^M (Graph 3, Figure 3) increased dramatically in unison responses from sessions #1-#3, followed by a decrease from sessions #4-#7. The number of unison responses peaked at session #8 (110 unison responses compared to 19 unison responses in #7) then decreased and increased again. The other exception was R^S (Graph 7, Figure 4), whose initial decrease in unison responses was followed by a zigzag pattern of increases and decreases. Increases in unison responses at sessions #4, #7 and #8 were consistently related to the use of specific techniques, such as tapping fingers on the target objects ("I see cat") or nodding foreheads together ("up", "down", "ball", "kick ball").

The second change involved imitative, conversational and spontaneous responses. As the number of conversational and spontaneous responses increased, the number of imitative responses decreased. For example, A^M (Graph 1, Figure 3) showed a decrease in imitative responses from sessions #3-#8, as spontaneous responses increased. Similarly, in the case of V^S (Graph 8, Figure 4), imitative responses decreased as target phrases were produced spontaneously and in conversational speech. There was a steady increase in conversational responses until session #12. The relatively even distribution of responses at session #12 was in contrast to the huge gap between imitative responses and all other levels of response at session #1. In the case of T^M (Graph 4, Figure 3), imitative responses decreased from sessions #8-#11 while spontaneous and conversational responses increased. As was seen with the unison responses, this trend was found in both groups, with the exception of two children, one from each group (J^M and R^S).

Figure 3. Number of responses at each level of intervention for children in the melodic group

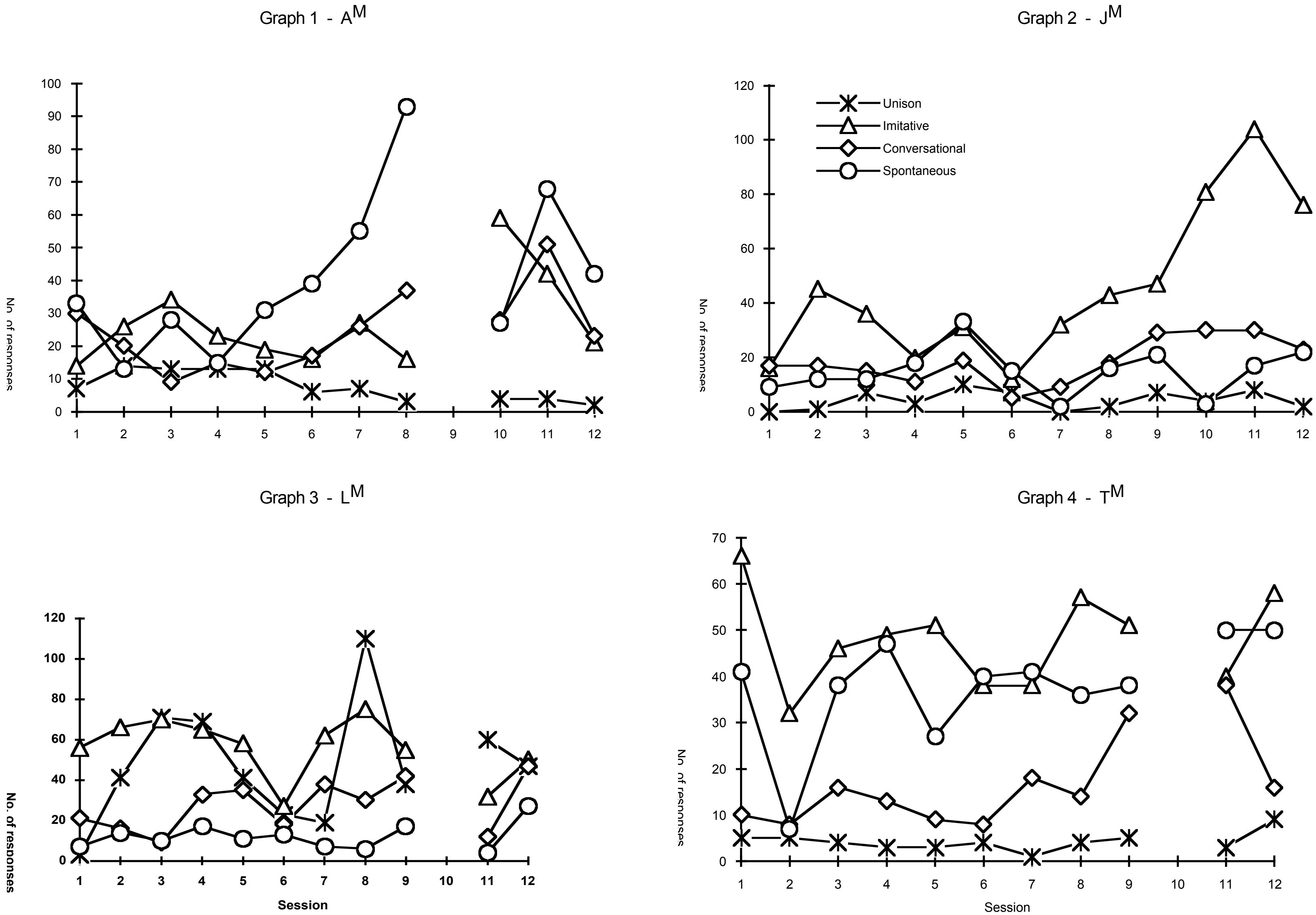
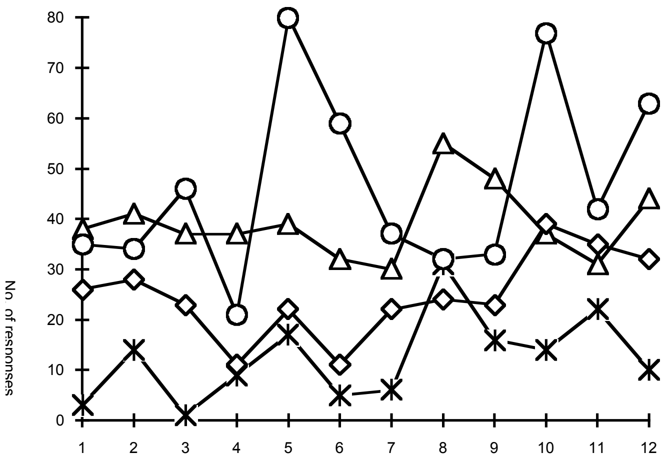
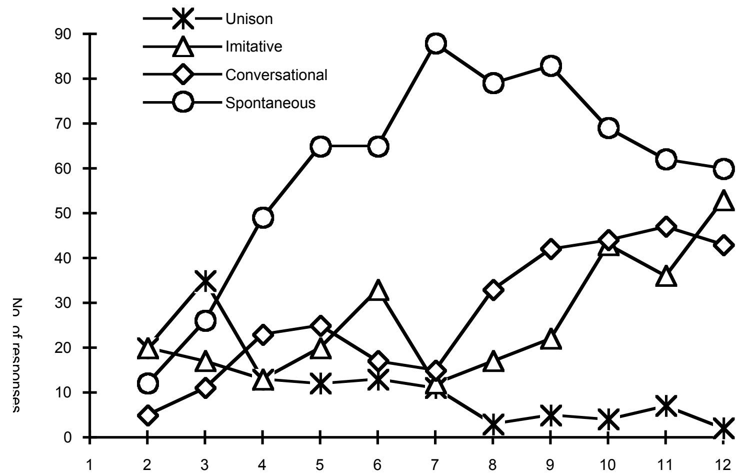


Figure 4. Number of responses at each level of intervention for children in the spoken group

Graph 5 - E^S

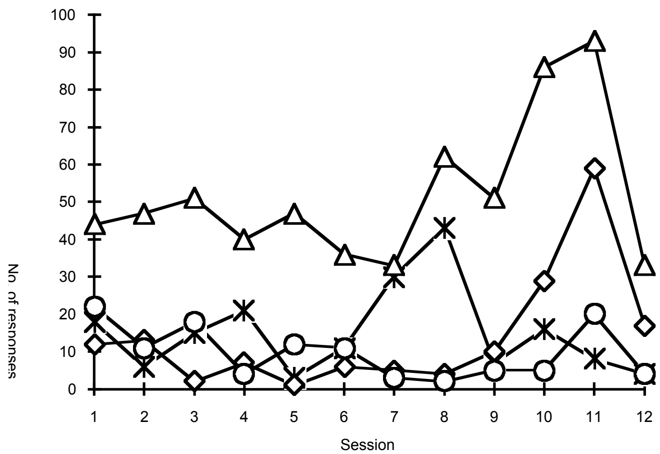


Graph 6 - J^S

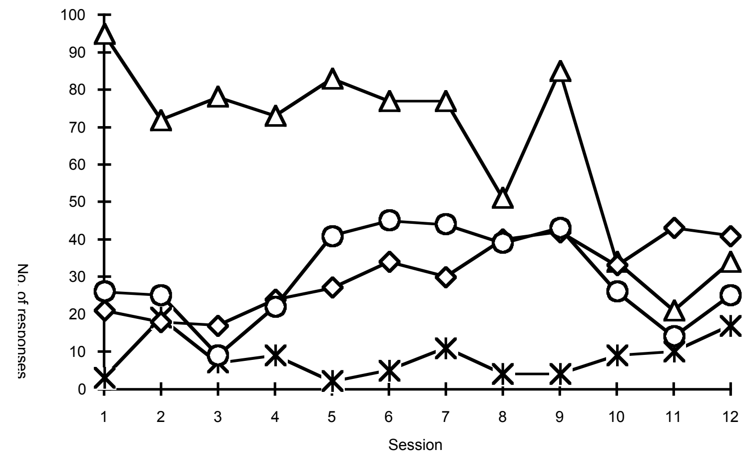


Graph 7 - R^S

Graph 8 - V^S



No. of responses
Session



No. of responses
Session

CLARITY OF SPEECH AND RATE OF RESPONSE

Fieldnotes revealed that regardless of group, children's verbal output became increasingly louder and clearer, with instances of unintelligible utterances gradually decreasing. In comparison, an increased rate or speed of response was particularly apparent in the melodic group, but observed infrequently in the spoken group. Children in the melodic group would respond to my queries more quickly and with fewer prompts or cues.

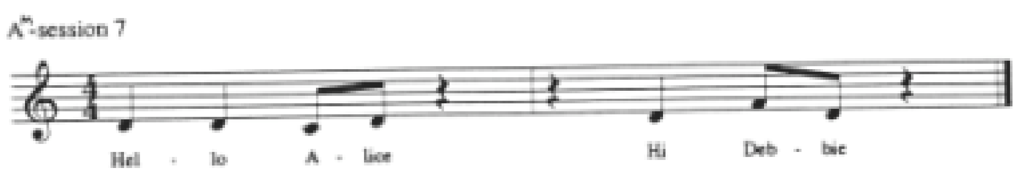
CONTEXTUAL FACTORS AFFECTING VERBAL OUTPUT

The following two excerpts were taken from fieldnotes midway through the intervention process.

EXCERPT 1: A^M

> A^M is looking out the dining room window as my car pulls up in front of the townhouse. When I arrive at the door, she and "baby" (a doll), are there to greet me. Her mother, who is in the basement, shouts "Hi". A^M takes my bag and carries it upstairs, leading the way to her parent's carpeted bedroom - a small, cluttered room with clothes and lots of knick knacks lying about. A^M heads for the only floor space at the foot of the bed and sits down. She takes the recorder out of the bag, plugs it into the wall, presses the "record" button. I sing "Hello ___" Without missing a beat, she responds "Hi Debbie". We repeat this musical greeting four times (see Figure 5) and with each repetition, her response is quicker and louder. Without wasting a moment, she takes the ball out of the bag. "What does baby want to do?", I ask. "Ball", replies A^M. I expand on her response and intone "Baby wants to play ball." Placing her mouth close to the recorder, she rhythmically chants, "Baby play ball please."

Figure 5. Musical greeting



EXCERPT 2: E^S

> I approach the playground of the hospital daycare. It is recess and E^S is running in the snow, a big smile on her face. When she sees me, she walks towards me and we continue on together down the path to the door. E^S heads to her locker. As she slowly takes off her hat, mitts, boots and snowsuit, she babbles unintelligibly (wanting to tell me something). Seven minutes after getting her from recess, we enter the small office, belonging to the director of the daycare. There is little space for playing in this room, which is dominated by a desk piled with papers, shelves, a filing cabinet and several chairs. E^S goes to the desk, picks up a pad of paper and babbles something about not touching anything on the desk - "no touch it" is the only understandable phrase. The voices of children just coming in from recess distract her. She goes to the door, holds the knob and babbles, "I heh dah dee goh nee". I interpret it to mean something like "I have to go now" by the way she imitates the intonational contours and rhythmic stresses of the sentence. I ask her to sit down next to me. I have placed the figures on the floor. She approaches, but goes straight for the recorder, "Touch it", "no touch it", she repeats several times (stereotypic response whenever she feels the urge to press the "Stop" button on the recorder). Once again I try to engage her in play but this time she is off to the window. "Horsie, horsie", she says (referring to the RCMP officer on horseback who came to visit the daycare two weeks before). Approximately eight minutes after entering the room, E^S sits down next to me and puts the daddy and mommy in the car. She repeats my spoken models, including "Daddy sit" and "Mommy (in) car", but only after I have repeated them several times. I am somewhat distracted by the children's voices in the playroom nearby, however E^S does not seem to be bothered and is able to remain on-task for 5 minutes at a time.

The differences reflected in these narratives suggest three key factors influenced the children's verbal output: the physical setting, tape recorder and manner in which the target phrases were presented (intoned vs. spoken).

A^M was at home waiting for me to arrive. E^S was at daycare playing outside with other children when I picked her up. Settling down to work in the director's office took eight minutes for E^S compared to no time at all for A^M. While the objects in the office stimulated E^S, they were a distraction, prompting unintelligible speech-babble and delaying on-task behaviour, particularly at the beginning of each session. The voices of children coming in from recess provided another distraction, resulting in less time spent producing intelligible verbal responses.

As for the tape recorder, it seemed to elicit quite different responses. A^M used it as a microphone, speaking loudly and clearly into it (telling me and "baby" what to do). In the session, from which the above excerpt was taken, 149 responses were recorded compared to 115 in the previous one. For E^S, the presence of the tape recorder seemed to recall previous negative experiences in the way it consistently elicited the stereotypic utterance "no touch it".

With regard to the manner in which the target phrases were presented, A^M was in the melodic group and E^S was in the spoken group. The musical greeting that evolved focused her attention and elicited a quick and confident response. A^M did not miss a beat when greeting me ("Hi Debbie"). In comparison, my requests to E^S to sit down did not immediately grab her attention. She remained unfocused, distracted and inclined to speech-babble for the first ten minutes of the session. When she finally sat down, several models of the target were required before she would respond.

While specifics were not the same, the setting, tape recorder and manner in which the target phrases were presented were factors that were present in all the sessions for A^M and E^S. These factors were also at work for the other children.

SETTING

It was found that the location of the room within the larger setting (home, school or daycare) was an important factor affecting output. Three children, A^M, T^M and J^S were at home napping or playing alone (or with an adult - mother, father or babysitter) just before I arrived. For these children, settling down to work typically required less than two minutes. In contrast, the other five children - E^S, J^S, V^S, J^M and L^M were at daycare or school, and were members of a peer group a fair distance from where the sessions took place. In the case of L^M and V^S, sessions took place in a basement office. Getting there from their respective classes took close to ten minutes. It involved walking down a long hallway to stairs (at this point, L^M would peek into each classroom, distracted by the novelty of walking in the halls while other children were in class), descending them, then walking along another corridor past several rooms. Once in the office, L^M and V^S would fidget and behave in a noncompliant manner for at least three minutes before settling down and responding verbally in a consistent manner. I noted in my fieldnotes that the disruption in their respective routines and the time required to get to the room (as well as its relative novelty) seemed to have a negative effect on on-task behaviour. As with E^S, this was especially apparent at the beginning of each session when it typically took these children five to eight minutes to settle down. For example, when I picked up V^S for session #10, she was sitting on a rug in a corner of the classroom, looking at a class photo album. This interruption seemed to affect her behaviour during the session in that she was unconcentrated and wanted to leave in the middle. Verbal output dropped to 102 responses from 174 responses in the previous session. For R^S and J^M, who attended a daycare, they had to walk upstairs to the room where the sessions took place. Whereas R^S happily left what he was doing (i.e. puzzle, playing in gym) to come with me, J^M had difficulty making the transition, particularly in the first five sessions.

For three of the children, E^S, R^S and J^M, changing rooms in which the sessions took place also affected verbal output. In the case of E^S, whereas the change to a large classroom in session #4 produced less verbal output (78 responses compared to 107 and 158 in sessions #3 and #5 respectively), she was more attentive (i.e. only one model of the target was necessary in contrast to previous sessions) and less likely to speech-babble. For J^M and R^S, a change in setting from daycare to home in session #11 resulted in a dramatic increase in total verbal output for that session - 159 responses for J^M (compared to 118 in #10) and 180 responses for R^S (compared to 136 in #10 and 58 in #12). In session #12, there was an unexpected room change to an open area near the front door of the daycare. This resulted in no verbal output at all from R^S for the first ten minutes of the session (see Table 6).

Table 6
<u>Setting in which sessions took place</u>

	Home	Daycare	School
A^M	#1-5 sister's room #6-12 parent's room		
J^M	#11 den	#1-12 except #11 upstairs playroom	
L^M			#1-12 basement office
T^M	#1-6 den #7-12 his room		
E^S		#1-12 except #4 director's office #4 classroom	
J^S	#1-12 playroom		
R^S	#11 den	#1-10 upstairs playroom #12 open gym area near entrance	
V^S			#1-12 basement office

TAPE RECORDER

The tape recorder played a notable role in eliciting clear and self-assured verbal responses, thereby assuming functions (i.e. as a microphone) apart from its originally intended one, which was to tape the sessions for the purposes of analysis. For example, V^S often requested to hear an excerpt from the session, not only listening with heightened concentration, but responding to all of the target models, even the ones to which she did not respond during the actual recording of the excerpt! This phenomenon was observed in most of the children. As seen in the previous excerpt, E^S was the case where the tape recorder had a negative impact on verbal output.

MELODIC VERSUS SPOKEN PRESENTATION OF SPEECH PATTERNS

The third and only factor to effect a group difference was the manner in which the target phrases and questions were presented. As was the case with E^S in the earlier excerpt, a certain active tension was lacking in the spoken group between the spoken model and the child's verbal response, resulting in a reduced speed of response. For V^S, the spoken model did not seem adequate enough in focusing her attention or of reducing instances of the stereotypic utterance, "mamamama", which she seemed to use in order to block out my requests, causing a considerable delay in her response. Eliciting a response seemed far more effective when the target phrases were intoned. It was as if the intoned phrase invited or beckoned a response. The following narrative is a series of reflections that I wrote in my fieldnotes for the children in the melodic group - L^M, J^M, A^M and T^M :

> The melodic motif seems to serve as a vehicle for the speech pattern, helping it to move forward in time and space, "carrying" it through to the end, and lending a particular creative tension to each play context. It is as if the melodic intoning of a word or a phrase within given parameters - melody, intensity, rhythm and tempo - creates a shared musical space for the child-adult interaction that allows for playing, risking, and exploring within the "confines" of the rules.

In addition to eliciting speedier responses for all the children in the melodic group, the intoned presentation also stimulated the playful experimentation and transformation of the target models. Fieldnotes and session data forms revealed that all four children in the melodic group often modified the target phrases by extending them, changing a word (noun or verb), exploiting the musical elements of pitch and dynamics, even creating a new melodic motif (see Figure 6). In contrast, only one child, J^S in the spoken group, modified a target phrase. The model "Boy brushing teeth" became "**I see little** boy brushing teeth" in session #8. Table 7 shows the sessions in which modifications occurred.

Table 7
<u>Modification of target phrases during sessions</u>

		Sessions 1	2	3	4	5	6	7	8	9	10	11	12
Melodic	A^M	x	x		x	x	x	x	x	x	x		
	J^M	x		x	x		x	x			x		

Group	L^M	x x x x x
	T^M	x x x x x x
Spoken Group	E^S	
	J^S	x
	R^S	
	V^S	

N.B. "x" indicates that child modified target phrase at least once during the session.

Figure 6. Modification of target word or phrase

The melodically intoned presentation of the target phrase also provided a structure within which the child and researcher interacted in a pleasurable and expressive way. This was particularly significant for L^{M}, whose difficulty in reproducing certain sounds could have resulted in considerable frustration. Instead, he tried hard to repeat words and complete phrases clearly and in tune. As with all the children in the melodic group, L^{M} enjoyed sound-imitation games and experimenting with the pitch, intensity and duration of a word (see Figure 7). He seemed to intuitively sense that I was enjoying the sounds we were making as much as he was. During such exchanges, L^{M} did not produce any stereotypic utterances (ex. "no", "no way").

Figure 7. Exploring pitch and intensity

For J^{M}, the melodic intoning of phrases was effective in easing him into a shared space that was non-threatening, novel and intrinsically motivating. Establishing a rapport with him in this way was particularly significant, in light of his mistrust of females in general. According to his Mom, he associated all females with the nurses who had "hurt" him in the past.

EXCERPT FROM SESSION #5: J^{M}

> J^{M} is reluctant to come with me, so G. (a child care worker) offers to carry him upstairs. When we arrive at the room, I take J^{M} and G. leaves. J^{M} looks fearful and anxious and seems to be holding back tears. Holding him in my arms, I walk towards the ten posters hanging on the wall. I improvise a song about each picture. J^{M} is attentive and begins to hum along. When I finish singing, he seems relaxed and so I put him down next to the playmobile figures of a boy and a mommy and two "sleeping bags". He puts the figures to bed and then tells them to wake up and eat a snack. This play

context elicits 1 and 2-word phrases such as "boy", "mommy sleep" and "eat orange". Several minutes later, when I intone "ball", J^M matches my pitch, intoning "ball", then he continues to intone the word six times to a motif he creates on the spot (Figure 6). After identifying pictures with the drum and reading a book, I put "Peter Parrot" (hand puppet) on, who praises J^M for his hard work and intones "Bye". J^M responds, "Ba-Bye", and continues → "ba-tu" (= "thank-you"), developing this into a jazzy motif "tutudatu" which I repeat (Figure 8).

Figure 8. Call-and-response improvisation

J^M's anxiety and minimal output at the beginning of the session was in contrast to his playfulness and vocal expressivity during the closing improvisation. The following three sessions began similarly and with each passing week, there was less "warming up" time required and increased verbal output (39 responses in session #6 compared to 159 responses in #11).

In addition to illustrating the positive effect of our melodically intoned exchanges, the above narrative highlights two other factors that affected verbal output for all children, regardless of group, child-researcher relationship and play routines.

CHILD-RESEARCHER RELATIONSHIP

I observed that the children's growing familiarity with me resulted in growing assertiveness, playful teasing and increased motivation to communicate. In the case of T^{M}, he called me back to his room after session #7 was over, using 3 consecutive 2-word utterances ,"Ah man, come here, no fair", complaining that I was leaving. In earlier sessions, he seemed to withdraw and become sad, not saying a word. J^{S} showed his growing assertiveness in session #8 when he expressed his desire to label the objects and say the 3-word target phrases, something he had never done before in previous sessions. Similarly, in session #9, he rejected my suggestion to look at pictures, using words (i.e. "not this, no cat, book"), where once he would only have pointed to the object he wanted to play with, occasionally saying "this". This change was reflected numerically in total number of responses - 132 in session #8 compared to 152 in session #9). A^{M} responded loudly with a smile, "It's a bunny", in session #3, when I intoned "It's a pig". In session #7, acting silly, she blew into the microphone and whispered "caca", instead of intoning "hi Debbie!", as requested (84 responses in session #3 compared to 115 responses in session #7).

PLAYROUTINES

The above narrative from J^{M}'s fifth session described a core set of play activities that was typically used with all children, regardless of group. J^{M} enjoyed playing with the figures, reading a book and saying "bye" to Peter Parrot. Each of these play contexts elicited consistent verbal output, and the more interested and involved J^{M} was in the play situation, the more inclined he was to speak. This was typical of all children with one exception, T^{M}. When playing with objects, such as a car (during the first home visit) or a plastic horse (session #4), T^{M} was drawn into his own world to the exclusion of any verbal communication. As with J^{M}, J^{S} enjoyed playing with the figures, but for him, pretend play was more than just physically manipulating the figures, it was also giving them a voice:

> J^{S} makes the mommy, daddy, girl and boy "walk" up to the top of the "mountain" (J^{S} has turned the container upside down). They are hungry for a snack. J^{S} picks up the mommy and in a high pitched voice says, "I like banana". "Okay, eat banana", he responds in his normal voice (session #7).

Creating different scenarios with the figures (i.e. taking a walk, snacking, sleeping, taking a drive, etc.) sustained the interest of all the children and resulted in increased verbal output, regardless of group. In the case of E^S, most of session #10 was spent playing with the figures, and it was in this session that her total verbal output peaked at 167 responses (compared to 120 in #9 and 130 in #11).

According to my fieldnotes, hiding and finding an object was another play routine that consistently elicited high interest and verbal output. Although the manner in which this game unfolded was unique for every child, as with most of the play contexts, it involved a ritualised sequence of target phrases that increased in length as the sessions progressed. In the case of V^S (session #3), I asked her what I should do with the ball. "Boot", she said, and put the ball in my boot. I modelled "Ball in boot". She repeated each word separately. Then I pretended to look for it ("where's ball?", "under chair?", "in pocket?"). V^S retrieved it from the boot. I modelled "Here ball", and then another hiding place was found and the game began anew. V^S responded well to the predictability (repetition of targets) and novelty (new and unexpected hiding places) of this game, and by session #7, she began to repeat target phrases, such as "Where ball ?", "Here ball", "Get ball" in their entirety. As for J^S, he liked to hide the ball underneath him and give false clues as to where I might look. This play routine facilitated the production of up to 4-word phrases (i.e. "Debbie, look in box").

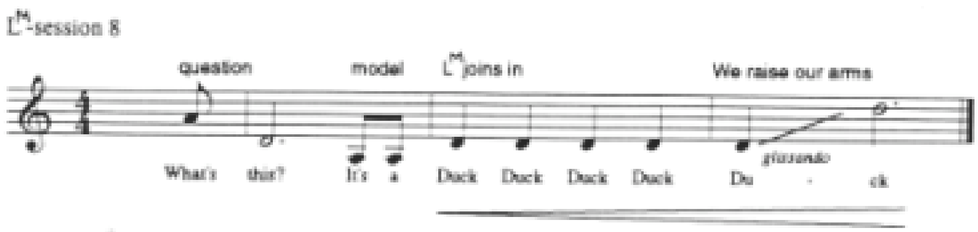
It was found that, within the context of particular play routines (i.e. identifying pictures or "reading" a book) the drum was a key factor in eliciting the different types of response: Unison, Imitative, Conversational and Spontaneous.

Certain rhythmic techniques with the drum (or in certain instances, a body part, such as the pointer finger or forehead) were particularly effective in eliciting unison responses. For E^S (Graph 5, Figure 4), the rhythmic hitting of the drum, as words and phrases were explored for the first time in session #8 (i.e. "sheep sheep sheep" ; "toast and but-ter") elicited the highest number of unison responses. As for J^M and V^S (Graph 3, Figure 3; Graph 8, Figure 4), tapping with our pointer fingers on a picture or a puzzle piece in rhythm to the target word or phrase (i.e. "dog"→ "dog here") elicited the highest number of unison responses in sessions #3 and #2 respectively. For R^S (Graph 7, Figure 4), tapping fingers on pictures ("I see cat") and nodding foreheads together ("up", "down", "ball", "kick ball") was effective in eliciting a high number of unison responses in sessions #4, #7 and #8.

In the case of L^M (Graph 3, Figure 3), there was an increase in unison responses in session

#3 and particularly at session #8, when I played with the duration (holding sound), rhythm (hitting the drum), intensity (getting louder or softer) and pitch (slowly sliding up five notes {interval of a fifth} or eighth notes {interval of an octave}) of each word or syllable. For example, in session #3, I held the first vowel sound of a word (i.e."du_____" for duck), waiting for L^M to join in, at which point we got louder or softer with each repeat of the word. This sequence was extended in session #8 to include a gradual increase in pitch and intensity, as we slowly raised our drumsticks until they were over our heads (Figure 9).

Figure 9. Technique used to elicit unison response



The drum was also effective in eliciting imitative responses in the way in which it established a physical boundary and encouraged turntaking. For example, in session #3, A^M and I each held a dolly, making it play the drum, as I modelled the target "I play drum". At the refrain ("LA lala LA LA"), we raised our dolls and made them dance. A^M then repeated the rhythmic pattern of the target. By session #7, she was able to repeat the intoned phrase in its entirety. As with A^M, most children, regardless of group, began to imitate increasingly longer phrases when the drum was part of the play experience. J^S spoke his first 7-word phrase in session #11 when the drum was used; J^M and E^S intoned or spoke their first 3-word phrase with the help of the drum in sessions #9 and #11 respectively.

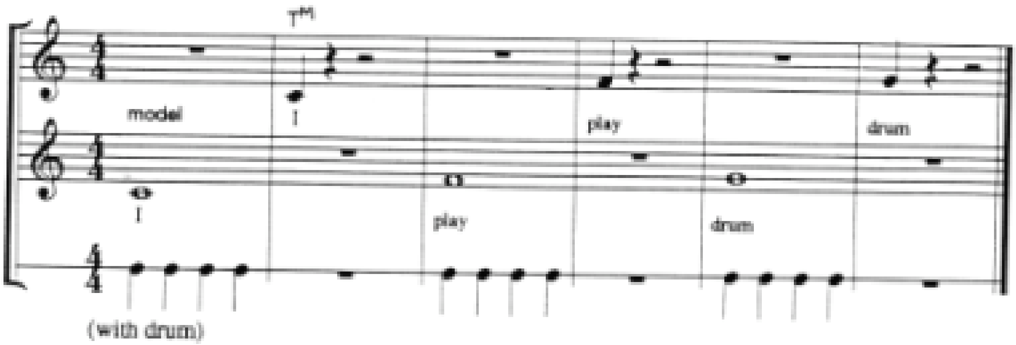
Using the same technique as I did with A^M for the phrase "I play drum" allowed me to observe the difference in the imitative responses of two other children, R^S and T^M, from the spoken and melodic group respectively. The technique consisted of using 4, 2 and finally 1 drum beat to reinforce each word. R^S was not able to go beyond uttering "Ah" after I modelled each word separately (Figure 10).

Figure 10. Technique used to elicit “I play drum” with R^S

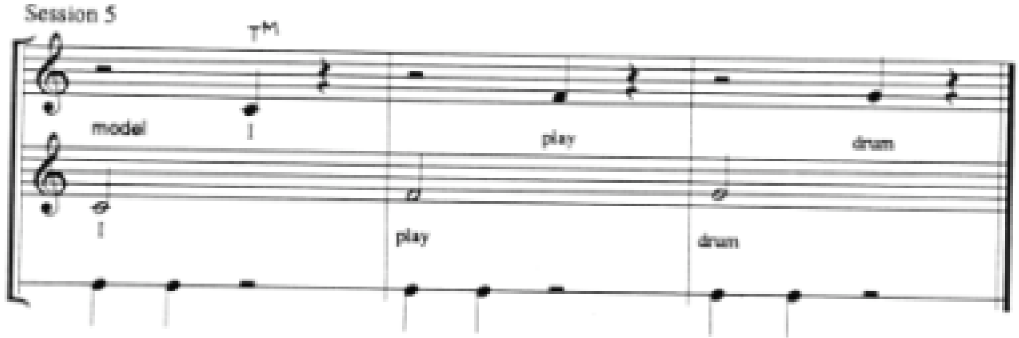
In contrast, by session #7, T^M succeeded in repeating "play drum" loud and clear with the drum, by first repeating each word separately after 4 beats, then 2 and finally immediately after the model (Figure 11).

Figure 11. Technique used to elicit "I play drum" with T^M

Session 3



Session 5

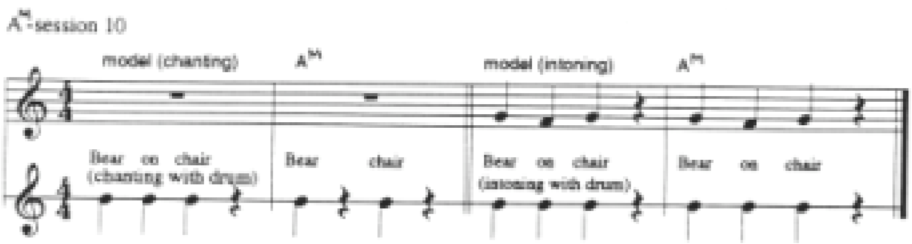


Session 6

Session 7

Similarly, with A^M (session #10), I noted a striking difference in her imitative responses when I first chanted, then intoned the phrase "Bear on chair" with the drum. Chanting the phrase elicited a hesitant, two-word response, whereas intoning it drew a speedier response that was louder, clearer and longer in length (Figure 12). Singing each word on a different pitch and drawing attention to certain ones by increasing intensity or pitch seemed to be effective in improving enunciation and increasing phrase length.

Figure 12. Chanting versus intoning “Bear on chair”



Furthermore, I noted in my session notes that the combined rhythmic and melodic stimulus also helped to reduce the tendency to run words into each other (slurred articulation). For example, by session #11, T^M intoned "move over" clearly and in tune, instead of "mohver". In comparison, the rhythmic element alone did not seem to be effective in the case of R^S, who, by session #10, was still saying "mumitting" instead of "mommy sitting", and "put kaway" instead of "put stick away".

In the case of T^M, the combined effect of the melodically intoned target and the rhythmic support of the drum resulted in his moving beyond merely imitating the intonational contour of my question to adding an answer onto it (Figure 13). A basic dialogue process was set in motion for T^M as imitative speech gradually began appearing in conversational speech.

Figure 13. From imitative to conversational speeh

For two children from the spoken group, J^S and V^S, and one child from the melodic group, J^M, the unelicited repetition of the last word or words of a question was often observed. They seldom went beyond just mirroring what the other person said.

Data also revealed that the use of the drum and a puppet facilitated role playing and conversational responses, regardless of group. For example, I would pretend to be "Super Bunny", "Peter Parrot" or Dolly, and ask the children questions, using the drum as a rhythmic support for the speech pattern. This technique was almost always successful in eliciting the desired response.

The drum was also an important factor in accounting for the inverse relation that was found between imitative responses and spontaneous responses. In both groups, it was effective in stabilising the use of longer target phrases, as well as in facilitating the spontaneous (internalised) use of these phrases within the context of play routines, where the child played the adult role:

J^S - SESSION #11:

> Today, J^S is the "teacher". He places the drum and book close to him, making sure that I can see all the pictures. He models a 3-word target phrase (ex. "I see flower"). His rhythmic drum support is very articulated as is the clarity of each word. When I repeat the phrase, he nods his approval. We continue in this manner until the book is finished.

By the end of the intervention period, five of the eight children, J^S, R^S, A^M, J^M and T^M, were using target phrases spontaneously in self-initiated play contexts using the drum. For example, supporting the speech rhythms on the drum, A^M chanted clearly, "I see cat, you do it". I repeated the phrase and we continued in this manner until all the pictures were identified (session #12). The remaining three children, in contrast, rarely initiated play contexts with or without the drum, and there were few to zero instances of the spontaneous supporting of speech rhythms without my modelling.

As the children moved from using the target phrase imitatively to producing it spontaneously (or in response to a question), it was noted that two children, J^S and R^S, generalised the role of the drum as a supporter of speech onto other objects. For example, finding it difficult to remove the lid of the playmobile container, J^S said, "Debbie help me", as he shook it in rhythm to the speech pattern (session #7). In the case of R^S, while looking at pictures in a book during session #11 (which took place at home), he used a plastic tabletop instead of the drum to support his speech.

Finally, for most of the children, including all four in the melodic group, the drum was effectively used in sound imitation games and word improvisations as a break from "work" during the session or to close the session.

PRE- AND POST-INTERVENTION DIFFERENCES IN VERBAL OUTPUT

Measures for the three dependent variables, mean length of utterance in morphemes (MLU), total number of words and production time (in minutes), were obtained from pre- and post-intervention language samples, each containing 100 utterances, produced by the child while interacting with a parent or myself. The mean length of utterance, or MLU, measured the child's level of syntactic development. The total number of words indicated the amount of verbal output by the child. Production time was expressed by the number of minutes it took to produce the language sample.

The scores for these dependent variables were then compared, using two statistical tests: a multivariate repeated measures analysis of variance (MANOVA) was calculated on each of the measures to determine whether post-intervention gains between groups differed significantly; the Pearson product-moment correlation coefficient measured the degree of association between pre- and post-intervention scores. Scores for each subject by group appear in Table 8. Means and standard deviations for each group appear in Table 9.

Table 8

Pre- and post-intervention scores for each subject

		Pre-intervention MLU	Pre-intervention Total number of words	Pre-intervention Production time	Post-intervention MLU	Post-intervention Total number of words	Post-intervention Production time
Melodic Group	A^M	1.59 .801 (sd)	163	14.5 min.	2.18 1.337(sd)	239	9 min.
	J^M	1.39 .598 (sd)	140	33 min.	1.44 .516 (sd)	148	15 min.
	L^M	1.12 .256 (sd)	116	14 min.	1.13 .230 (sd)	114	7.5 min.
	T^M	1.26 .335 (sd)	134	14 min.	1.69 .771 (sd)	179	6.5 min.
Spoken Group	E^S	1.37 .365 (sd)	147	23.5 min.	1.26 .610 (sd)	150	11 min.
	J^S	1.18 .296 (sd)	125	9 min.	1.46 .655 (sd)	146	7.5 min.
	R^S	1.13 .365 (sd)	134	22.5 min.	1.44 .589 (sd)	151	17 min.
	V^S	1.17 .203 (sd)	119	13 min.	1.22 .438 (sd)	124	15 min.

Table 9

Means and standard deviations for each group

		Melodic Group (N=4)		Spoken Group (N=4)		Entire Sample (N=8)	
		Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Total words	Pre	138.25	19.40	131.25	12.18	134.75	15.45
	Post	170.	53.11	142.75	12.69	156.38	38.60
MLU	Pre	1.34	.20	1.21	.11	1.28	.16
	Post	1.61	.44	1.345	.123	1.48	.33
Production time	Pre	21.13	9.10	14.75	5.66	17.94	7.80
	Post	10.00	3.85	12.13	4.59	11.06	4.08

The findings for each dependent measure were as follows:

TOTAL NUMBER OF WORDS - The difference between pre- and post-intervention scores for both groups was slightly significant (p =.057), as shown in Table 10. All but one child produced a greater number of words in the posttest than in the pretest.

Table 10
Total number of words for entire sample

Source	*SS*	*df*	*MS*	*F*
Within groups	2035.90	6	339.65	
Time	1870.56	1	1870.56	5.51 *
Group by time	410.06	1	410.06	.314

* p = .057

MLU - Pre-and post-intervention mean scores for the entire sample (N=8) ranged from 1.28 to 1.48. Although the difference between the two groups was not statistically significant (p =.060), as shown in Table 11, post-intervention gains for the melodic group (1.34 to 1.61) were somewhat greater numerically than for the spoken group (1.21 to 1.33). It was also shown that children with high incoming MLU scores tended to have high MLU scores at the end of the intervention. This was reflected in the significant correlation (.738; p<.05) that was found between pre-and post-intervention scores for MLU.

Table 11
Mean length of utterance (MLU) for entire sample

Source	*SS*	*df*	*MS*	*F*
Within groups	.18	6	.03	
Time	.16	1	.16	5.36 *
Group by time	410.06	1	.02	.63

* p = .060

PRODUCTION TIME - There was a significant decrease in the amount of time required to produce 100 consecutive utterances for both groups ($p < .05$; Table 12), however the effect was greater for the melodic group than for the spoken group ($p < .05$; Table 13, Figure 14). There was also a significant correlation (.994; $p < .01$) between pre- and post-intervention gains for the melodic group. As illustrated in Table 8, it took half the time for the children in the melodic group to produce 100 utterances in the post-intervention home visit than in the pre-intervention home visit.

Table 12
Production time for entire sample

Source	*SS*	*df*	*MS*	*F*
Within groups	61.44	6	10.24	
Time	189.06	1	189.06	18.46 *
Group by time	72.25	1	72.25	7.06 **

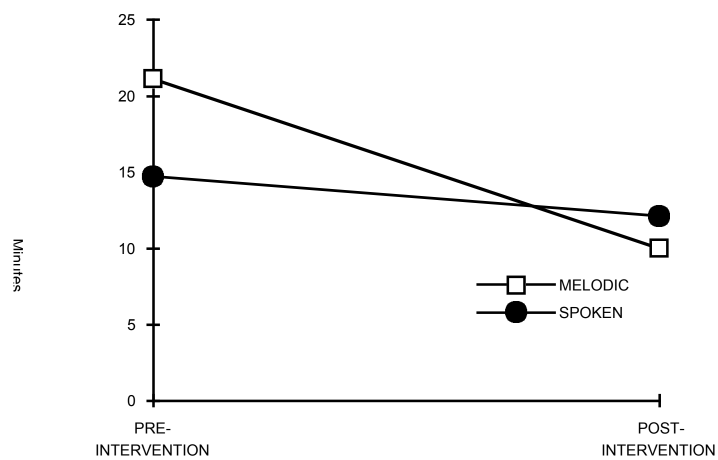
* $p<.05$ ($p=.005$) ** $p<.05$ ($p=.038$)

Table 13
Production time for the melodic group

Source	*SS*	*df*	*MS*	*F*
Within cells	41.84	3	13.95	
Time	247.53	1	247.53	17.75 *

* $p<.05$ ($p=.024$)

Figure 14. Pre- and post-intervention group means for production time



Finally, the coded transcripts of the pre- and post-intervention language samples revealed that there was a decrease in the number of unintelligible utterances for four of the children, regardless of group (Table 14).

Table 14
Number of unintelligible utterances from pre- and post-intervention language samples

		Pre-intervention sample	Post-intervention sample
Melodic Group	A^{M}	6	7
	J^{M}	13	5
	L^{M}	1	1
	T^{M}	14	7
Spoken Group	E^{S}	6	0
	J^{S}	5	6
	R^{S}	19	26
	V^{S}	27	6

CHAPTER V

SUMMARY, DISCUSSION, CONCLUSIONS, IMPLICATIONS

SUMMARY

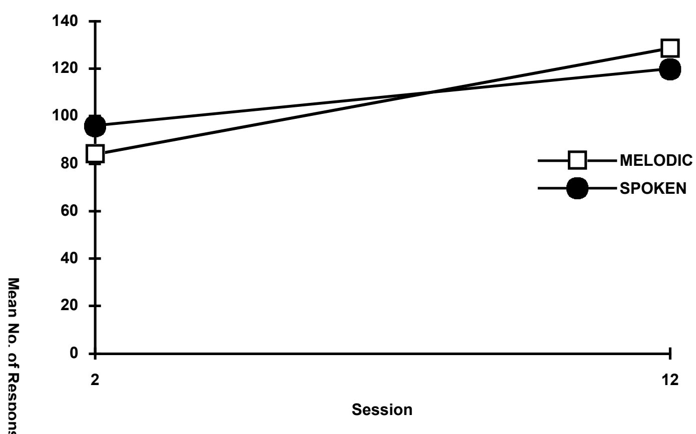
The present study examined the effectiveness of an adaptation of Melodic Intonation Therapy (MIT) in increasing the communicative speech of young children with Down syndrome. Eight children were matched on the basis of their mean length of utterance (MLU) and randomly assigned to one of two groups. Each child in the melodic group received twelve-weekly 30-minute individual sessions, carried out by myself. Each child in the spoken group received the same treatment in all respects except for the melodic component. More specifically, the only difference between the two groups was in the manner in which the target phrases, questions and cues were presented: they were spoken for children in the spoken group and intoned for those in the melodic group.

Data was collected from pre- and post-intervention home visits and from the twelve weekly intervention sessions. Scores for three dependent measures - total number of words, mean length of utterance and production time or rate of response (time required to produce 100 consecutive utterances), were obtained from the transcripts of pre- and post-intervention language samples of children at play, first with a parent and then with myself. Fieldnotes were written, following the two home visits and the weekly sessions, and examined for contextual factors, as they affected children's verbal output. Every verbal response during each session was noted and categorized according to the four levels of response: unison, imitative, conversational and spontaneous. The number of responses within each category as well as the total number of responses for each session were then computed.

Data from the intervention process showed that the only factor which effected a group difference was the melodic versus spoken manner in which speech patterns were presented, whereas interconnected contextual factors, such as the physical setting, child-researcher relationship and the play routine influenced verbal output, regardless of group.

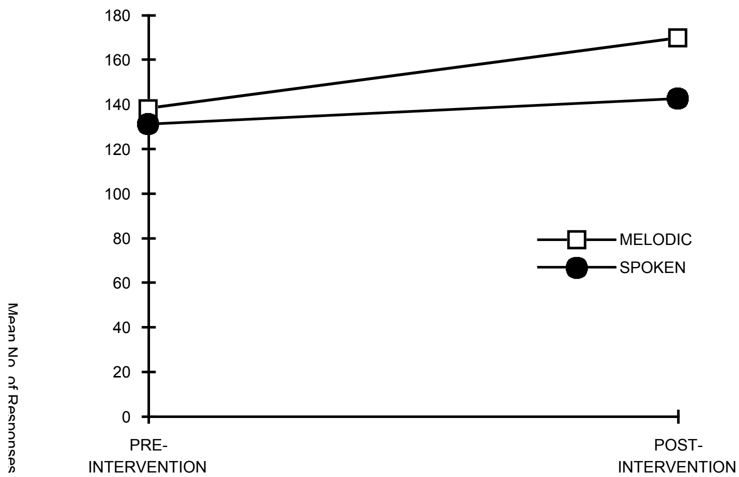
It was found that children in the melodic group were quicker to respond (increased rate of response) and experimented more with the target phrases by modifying, extending or transforming them. These observations were made less frequently with children in the spoken group. This group difference was expressed numerically by a slightly greater average increase in total number of responses for the melodic group than for the spoken group (see Figure 15).

Figure 15. Group means for total number of responses during intervention process



A comparison of the pre-and post-intervention scores for the total number of words and production time revealed similar differences between the melodic and spoken groups. Whereas there was a marginal effect for total number of words for both groups (p = .057), this effect was largely attributed to the pre- and post-intervention gains for the melodic group, which were greater than for the spoken group, as illustrated in Figure 16 (138-170 words for melodic group compared to 135-156 words for spoken group).

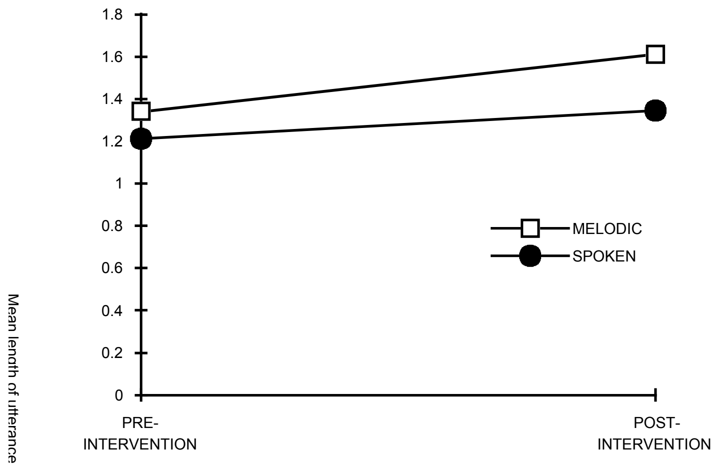
Figure 16. Pre- and post-intervention group means for total number of words



With regard to production time, although it took both groups significantly less time to produce 100 utterances in the post-intervention language sample ($p < .05$), children in the melodic group produced the utterances in a significantly shorter period than children in the spoken group, requiring half as much time than they did in the pre-intervention language sample (correlation coefficient of .994; $p < .01$).

As for MLU, while it was not measured each session, an indication that MLU was improving was reflected in the increases in range of length of utterance that was observed (Table 5). Two of the three children who improved the most were from the melodic group. Statistical data from the pre- and post-intervention language samples told a similar story. A marginally significant effect for MLU (p =.060) was found, which was almost entirely due to the post-intervention gains in the melodic group, as illustrated in Figure 17. As well, the significant correlation that was found between pre- and post-intervention scores for MLU indicates that the incoming MLU had an effect on the magnitude of the gains made.

Figure 17. Pre- and post-intervention group means for mean length of utterance (MLU)



DISCUSSION

MELODIC VERSUS SPOKEN PRESENTATION OF SPEECH PATTERNS

Findings from this study suggest that the MIT intervention was effective in increasing verbal output and rate of response. Session data revealed that children in the melodic group were quicker to respond than children in the spoken group. They also experimented more with the speech patterns by modifying and lengthening them. Similarly, post-intervention scores revealed that children in the melodic group produced more words in less time than children in the spoken group. These findings support previous reports of MIT's effectiveness in improving speech production (Sparks et al, 1974; Marshall & Holtzapple, 1976; Krauss & Galloway, 1982). As well, data from both the weekly sessions and the pre- and post-intervention language samples showed that whereas a gradual increase in the length of utterance was found for both groups, there was a slightly greater improvement in mean length of utterance for children in the melodic group than for those in the spoken group. From these findings, one can speculate that a larger group difference, or a delayed effect for the melodic group, might have been detected if language measures would have been taken several months after the end of treatment. Sparks et al (1974) found that syntactic growth began to appear post-MIT. Similarly, Miller and Toca (1979) described the case of a 3-year-old boy who began to show improvement in combining words one month post-MIT.

How can we account for these group differences in verbal output? Three possible explanations can be identified. Firstly, the melodic component might have added a dynamic dimension (musical dimension) to the play situations, possibly strengthening their influence on the children's total verbal output. I noted in my fieldnotes that the melodic component seemed to serve as a vehicle for the speech pattern, carrying it through to the end, and inviting a response. Intrinsic to the melodic pattern was a sense of structure and expression, which helped to create a playful space, allowing for pleasurable interactions and freedom in exploring the sounds of speech, much like the early pre-linguistic dialogues between mother and child.

Developmental researchers have recognised the central role that melodic intonation plays in the pre-linguistic communication experiences between parent and child (Leung, 1985; Fernald, 1989). The child "sings" long before s/he speaks. At first, the child expresses its needs through the intonational patterns of its crying and cooing (Fernald, 1989). By the third month, the child begins to explore vowel sounds and to discover the satisfaction of hearing its own voice and communicating with another person. These early pleasurable communicative experiences provide a base for further language development. When consonants are added to the sound repertoire at about six months, the child's babbled utterances begin to include nonverbal prosodic patterns of speech (Zoller, 1991). The first words appear by 12 months, followed by two-word utterances from 18 months.

A second explanation to account for the findings was the interplay of the drum and the intoned phrase. Data showed that when the drum was used to reinforce the rhythm of the intoned phrase (as was the case for the melodic group), there were group differences with respect to clarity and rate or speed of response. There were also numerous instances of the modification and transformation of target phrases. If one considers the rhythmic element as a cohesive force, highlighting and organising patterns (left hemispheric functions) and the melodic element as an expressive force, a source of emotional satisfaction (right hemispheric functions), then one might appreciate the bilateral stimulation and increased attention and motivation, resulting from the combined effect of drum and intoned utterance.

A third explanation to account for group differences in verbal output was that I enjoyed the experience of intoning phrases more than speaking them, feeling less restrained when interacting with the children in this way. This raises the issue of determining to what degree one can separate the effects of a particular intervention, in this case, MIT, from the perceptions and experience of the clinician, who is interacting with the child in what Bunt (1994) described as "a pleasurable joint activity".

Music, particularly its prosodic elements, speaks to the emotions. Damasio and Damasio (1977) distinguished the sort of verbal language processing, which enables one to sing the lyrics of a song, from the processing necessary for uttering the same words outside a musical context. The latter process involves a more analytical construction and is generated by left hemisphere function, while the former process is most probably generated by the right hemisphere and is closely related to emotional experience and expression. The pleasure and emotional satisfaction in actively engaging in a music-based experience cannot be overlooked.

CONTEXTUAL AND SUBJECT FACTORS AFFECTING VERBAL OUTPUT

Intervention process

Analysis of the session fieldnotes revealed that, within the larger social-interactive play context in which the intervention process unfolded, there were certain interconnected factors affecting the children's verbal and intoned responses, regardless of group. It was noted that the physical setting, the child-researcher relationship and the play routine affected total output (as expressed by the total number of responses). These findings support the view of developmental theorists and researchers, who have long recognised the important influences of the sociocultural and play contexts on speech development (Vygotsky, 1978; Conti-Ramsden & Snow, 1990). Language is acquired through the dynamic interactions with people and objects in the child's environment (McLean & Snyder, 1978). In the early years, the physical and the social world seem to intertwine much more closely than has been assumed, as the child communicates with another person about a "shared world" (Uzgiris, 1981) As a "facilitator" of language, the adult's role is to manipulate the child's physical, social and linguistic environments in order to stimulate language (Bloom & Lahey, 1978; Vygotsky, 1986).

The role of play has also been considered by developmental theorists including Piaget and Vygotsky. Piaget's belief that the frequency of speech is in proportion to that of imaginative play (1959) is consistent with present findings that children's verbal output increased while creating different scenarios with playmobile figures, such as taking them for a walk, giving them a snack or putting them to bed. In Vygotsky's view, play with another person provides the child with the first opportunities for social, cognitive and linguistic growth. He wrote, "As in the focus of a magnifying glass, play contains all developmental tendencies and is itself a major source of development" (Vygotsky, 1978, p.102).

As for the physical setting, it was found, that for those children who were seen at school or daycare, taking them from a peer group context to another room some distance away, had a negative effect on on-task behaviour and subsequently verbal output. Bryant and Graham (1993) noted that there is no evidence to suggest that individual therapy is superior to group therapy, and that furthermore there is possible stigmatization by peers, that is, children that are being taken out of class may be labelled as stupid or backward by their classmates.

There were two other notable findings. Data revealed that most children, regardless of group, showed an increase in the length and clarity of response, particularly when the drum was used as part of the play routine to support the rhythm of the speech patterns. Evidence of the effect of the drum was also seen in the pre- and post-intervention language samples, where the number of unintelligible utterances decreased for four of the children, regardless of group (see Table 14).

The drum also influenced change within the different levels of response (Unison, Imitative, Conversational and Spontaneous) and might have accounted for the inverse relation that was found between imitative responses and conversational and spontaneous responses. These findings provide further evidence of the important role of imitation in speech development (Tudge, 1990); more specifically, words that were once imitative would gradually come to be used spontaneously (Bloom et al, 1974). Scarpa (1990) underlined the importance of imitation and repetition in the development of conversational skills. The move from imitative to conversational speech was seen with one child, T^M, during the intervention period. This basic dialogue process was described by Scarpa as the interplay between mirroring what the other person says (specularity) and adding something on (complimentarity).

Finally, it was found that, by assuming a function (i.e. as a microphone) apart from its originally intended one, the tape recorder facilitated articulate responses. This did not happen, however, during the collection of the pre- and post-intervention language samples, where the tape recorder went largely unnoticed.

CONTEXTUAL AND SUBJECT FACTORS AFFECTING VERBAL OUTPUT

Collection of pre- and post-intervention language samples

Examination of the context in which the language samples were collected allowed me to appreciate to what extent statistical data was limited in its ability to give the full picture, and to what extent variability in the children's scores was a result of various subject factors (often at work at the same time) affecting verbal output. These included sociocultural influences, such as parental attitude and interactional style. For example, it was illustrated how a primarily directive approach limited the child's role to that of a follower, whereas a balance between a non-directive and a directive approach allowed for the child's participation as a partner. In the case of L^M, the primarily directive approach of his father limited L^M's verbal output which was reflected in the pre-intervention language scores. Children's medical conditions might also have affected verbal output: two children, J^M and T^M, were undergoing audiological testing at the time of the study; three children, J^M, T^M and J^S had tubes inserted in their ears to reduce fluid build-up [8]; two children, J^M and V^S wore hearing aids, but not all the time; one child, L^M, had speech motor sequencing problems (also known as expressive apraxia). Finally, the child's emotional and physical state (determining variables such as sickness, time of day, etc.) affected verbal output at the time of the language sampling. In the case of J^S, his mother's illness during the post-intervention language sampling appeared to have had an inhibiting effect on his verbal output. Whereas his total verbal output increased from 57-158 responses over the course of the twelve-weekly sessions, this increase was not at all reflected in the outcome scores. As for T^M, he had a sinus and ear infection

[8] Middle ear effusion or otitis media can lead to conductive hearing loss.

during the pre-intervention language sampling and this might have affected language scores, particularly with regard to the magnitude of the gains made.

CONCLUSIONS

Based on findings from this study, it can be concluded that:

1. The Melodic Intonation Therapy (MIT) intervention had an effect on the children's total verbal output and rate of response. It also encouraged the experimentation of the target phrases.

2. MIT was an effective method for stimulating verbal speech in the way it mirrored early language development by exploiting the prosodic characteristics of speech.

3. The interplay of the drum and the intoned phrase might have accounted for the marginal differences between the melodic and spoken groups with regard to MLU and clarity of production (speech intelligibility).

4. Children's incoming mean length of utterance (MLU) had an effect on the magnitude of the gains made; children with high incoming MLU tended to have high post-intervention scores and the opposite was also true. All the subjects in this study had incoming MLU of 1.6 or less, which might have accounted for the gradual improvement that was observed.

5. Contextual factors, such as the physical setting, child-researcher relationship and play routine affected children's verbal output during the intervention sessions, regardless of group.

6. The drum played a key role for all children in increasing the length and clarity of response. It was also an important factor in effecting change within the levels of intervention, particularly with regard to the inverse relation that was found between imitative speech and

conversational and spontaneous speech.

IMPLICATIONS FOR FUTURE RESEARCH

Present findings of a significant correlation between pre- and post-intervention MLU measures are consistent with those of Sparks et al, 1974, who found that subjects, who made the greatest gains in phrase length, produced some stereotyped jargon at the outset of treatment, and subjects, who made moderate gains, had produced little or no verbal output. These results have implications for establishing criteria for MIT candidacy in young children. For example, if incoming MLU is less than 1.5, a longer treatment period would be needed in order to determine more precisely the impact of MIT on the syntactical development of young children with Down syndrome, particularly in light of child language data that suggest that the move from one-word to two-word utterances can take a relatively long time (Brown, 1973). Children with a higher incoming MLU (1.5-2 or more) might benefit more from Melodic Intonation Therapy with respect to syntactic growth. Helfrich-Miller (1980) suggested that the child with an MLU of 3 or 4 would be a good candidate for MIT. In order to establish the durability of gains made, language samples might be taken during and after the intervention at regular intervals.

It would also be important to identify more precisely the contextual factors at work during the intervention process and during the collection of pre-and post-intervention language samples (i.e. play context, evolving child-researcher relationship, physical setting, subject differences) and how these factors potentially affect verbal output. This study has exposed some of the dangers of using spontaneous language sampling as a measure of verbal competence as well as the need for understanding the various contextual and subject factors affecting verbal output at any given moment. Perhaps, it would be important to match children on the basis of several variables, including MLU, total verbal output and oral-motor abilities, instead of only one variable such as MLU (i.e. L^M was the only child with expressive aphasia, which limited his capacity to combine not only words but also the syllables in a word.

With regard to the recording of the language sample, while audiovisual documentation would allow for the examination of other variables of expressive language (i.e. gestural expression), its very presence would be distracting and therefore a potential confounding factor. For example, one child might "act out", while another might become withdrawn in the presence of a video camera. In a study of autistic children and their mothers, Warwick (1988) found that the video camera hampered physical and psychological space. Parents felt more relaxed to react spontaneously without the video camera. In the present study, it was found that a small tape recorder was unobtrusive and easily hidden from sight, thus maximising the naturalness of the child-adult interactions during the two home visits. It was interesting to note, however, that during the intervention sessions, the tape recorder became a play object (i.e. microphone), influencing verbal output.

Finally, an aspect worthy of investigation, is the role of imitation in MIT, more specifically, the trend from imitative speech to conversational and spontaneous speech.

IMPLICATIONS FOR CLINICAL PRACTICE

Findings from this study have implications for clinicians, both music therapists and speech-language pathologists, for the implementation of Melodic Intonation Therapy with young children. Firstly, in determining when to begin MIT treatment, present data suggests that MIT is an effective method for facilitating the verbal output of young children, who are at least at the stage of uttering one word, and for promoting syntactic growth in children whose mean length of utterance is greater than 1.5. These results are supported by child language data that suggest a close relationship between the vocabulary spurt, occurring between 18 and 24 months, and the beginning of two-word utterances (Fowler, 1990).

As well, evidence of the potentially disruptive and delaying effect of removing children from their daily school or daycare routine in order to treat them individually, raises the issue of the effectiveness of "pull-out" programs. According to Bryant and Graham, there is a trend favouring in-class therapy programs. Three advantages have been stated: 1) the learning environment is more naturalistic, with normal children possibly serving as effective peer models (Humpal, 1990)[9], 2) the use of the therapist's time is cost-effective and 3) there is a chance to model and train staff. In light

---

[9] Byrant and Graham (1993) suggested that in a group of four children, two may be normal. However it would be necessary for them to have been involved in an integrated program for at least six months prior to the onset of treatment.

of the above-stated benefits, the implementation of inclusion programs, where the child works with peers within the context of the group, might warrant future consideration.

This study also revealed that the play routine was an important factor affecting verbal output, and that within the context of certain play situations, such as reading a book or role-playing with a puppet, the drum played a key role for all children in increasing the length and clarity of response, as well as in effecting change within the different levels of response. These findings have implications not only for the implementation of MIT with children, but also for the development of early language intervention strategies in general, particularly with regard to the choice of play materials (i.e. ball for physical play, playmobile figures and toy phone for imaginary play, hand puppets for role-play) and the use of a drum (or a body part used rhythmically - clapping hands, tapping index finger on target object, etc.) to reinforce the rhythm of the speech patterns.

Furthermore, findings that imitative responses gradually appeared in conversational and spontaneous speech underline the importance of imitation in language development and have implications for effectively choosing words and phrases[10] as the target of language stimulation. The most appropriate ones would be those that create a minimal discrepancy between what the child already knows and the next level of development - in Piagetian terms, phrases should be of "moderate novelty" (Bricker & Carlson, 1981); in Vygotskian terms, phrases should be in the zone of proximal development[11] in order to maximize potential (Tudge, 1980).

---

[10] For ideas on target phrases and play contexts in which to practice them, refer to "It Takes Two to Talk", by A. Manolson (1992, pp.62-28). For ideas on how to set speech patterns to music, refer to Hoshizaki (1983, ch.10, pp.90-95) and Marshall & Holtzapple (1976).

[11] Vygotsky termed the difference between the child's actual developmental level and immediate potential for development as the zone of proximal development (Tudge, 1990).

IMPLICATIONS FOR EDUCATIONAL PRACTICE

The present study also has implications and applications for educators and parents, particularly in light of the movement towards inclusive education [12] as a result of the passage of Section 15(1) of the Canadian Charter of Rights in April, 1985, assuring all children the right to a public education (Winzer, 1990).

The key to successful inclusion is an interdisciplinary approach to education with a focus on each child's abilities not disabilities (Darrow, 1990)

Music is a powerful integrating force in its ability to bring children of varying levels of functioning together in a fun and stimulating atmosphere. As a multisensory stimulus, music is easily accessible and can provoke different responses on many levels simultaneously (i.e. auditory, kinaesthetic, tactile, visual). Music-based experiences can stimulate and maintain the interest of the special needs student while offering opportunities for furthering the musical development of the group as a whole.

As a music-based language stimulation strategy, MIT can be a fun and effective way for children of varying abilities to work together. Within the context of a small group, children with speech delay can improve verbal communication skills, while their normal peers can help to reinforce the language learning that takes place. Furthermore, increased sensitivity to the melodic and rhythmic aspects of speech as a result of MIT (i.e. improved auditory discrimination, rhythmic imitation and vocal projection) can ensure the children's successful participation in integrated music classes, in particular, choral groups.

The notion that music may increase bilateral cerebral arousal levels, possibly through the mediating role of the right hemisphere (Morton et al, 1990) has particular implications for MIT's potential effectiveness in improving concentration, memory and on-task behaviour. Moreover,

---

[12] The notion of inclusion has been replacing the concept of mainstreaming, as it more accurately describes the need for all children to be included in all aspects of community and school life (Bryant & Graham, 1993)

MIT might be particularly effective with young speech-delayed children, who are not motivated to communicate because of repeated experiences of failure to meet the verbal expectations of those close to them (as was the case for L^M).

In order to maximize the effectiveness of the MIT intervention, clinicians, parents and educators might work together in determining the target phrases and preferred play situations. If the speech therapist is the primary therapist, it would be important to consult the music therapist regarding the melodic intonation of speech patterns. Once the phrases and play contexts have been established and treatment has begun, clinicians might involve the parents and teachers (i.e. first by modelling then by training them) in order to reinforce the language learning that is taking place. Previous studies underlined the need for parental input in order to consolidate gains (Marshall & Holtzapple, 1976; Sparks & Deck, 1994).

Dunst (1986) asserted that "we should no longer focus on the question of whether early intervention works but rather on how it works" (as reported in Thurman & Widerstrom, 1990). Before a better understanding of the effects of MIT on expressive verbal speech can be established, specific applications of the method must first be examined. By restricting the present investigation to children with Down syndrome, this study represented a first step in that direction. Future MIT research might help to determine the generalizability of the adaptation of MIT, as implemented in the present study, to other child populations with language delay and its practicability to an integrated (or inclusive) group instructional setting.

APPENDIXES

APPENDIX 1

FORMS AND LETTERS

LETTER SENT TO PROFESSIONALS AND ACTIVE PARENT MEMBERS OF ASSOCIATIONS FOR THE HANDICAPPED FOR FURTHER DISTRIBUTION

Dear Parents,

ARE YOU INTERESTED IN HAVING YOUR CHILD PARTICIPATE IN TWELVE LANGUAGE STIMULATION SESSIONS IN THE FALL?

I am looking for young children with Down syndrome (ages 2-8) to participate in a research study. This study will examine the effect of a language stimulation programme, known as Melodic Intonation Therapy, on the development of expressive speech in children with Down syndrome. The study will consist of twelve fun and stimulating sessions for your child.

Please contact me at before June 20th.

Sincerely,

Debbie Carroll
Graduate student, McGill University

LETTER SENT TO PARENTS WITH ENCLOSURES

September 19, 1993

Dear ,

Thank you for agreeing to have T. participate in my research project examining the effectiveness of a language stimulation method, known as MELODIC INTONATION THERAPY. This intervention has already been shown to be beneficial for children with speech delay.

The project will involve two home visits (one in September, the other in December for approximately 45 minutes each) to observe your child at free play. It will also consist of 12 weekly 25- minute language stimulation sessions for your child (time and place to be confirmed). During these sessions, your child will be asked to imitate and initiate speech patterns that are consistent with his current level of speech development. Puppets, pictures and body actions will be used in fun and pleasurable ways to reinforce the meaning of the speech patterns.

In order to effectively plan these sessions, I will need some information regarding your child's present level of language production and comprehension. Kindly fill out the enclosed language profile form, checklist and short questionnaire by the first home visit on September 27th.

Please be assured that any information that you provide and any data collected during the study will be held in strict confidence. At no time will your child's name be mentioned. You will be free to withdraw your child at any time during the study.

Thank you for your cooperation. I look forward to meeting you and T. on Monday, September 27 at 5:45 p.m.

Sincerely,

Debbie Carroll,
Graduate Student,
McGill University

INFORMED CONSENT

I acknowledge that I have been informed of, and understand the nature, purpose and procedures of this study, and I freely consent to have my child participate. I have also been informed of my right to withdraw my child at any time during the study.

Date ____________________________
Signature of parent_____________________________________
Telephone number_____________________________________

LANGUAGE PROFILE FORM

(adapted from "It Takes Two to Talk" by A. Manolson, 1992)

NAME OF CHILD :
NAME OF PARENT :
DATE :

PLEASE USE THE CHART BELOW TO RECORD YOUR CHILD'S EFFORTS TO COMMUNICATE (GESTURES, SOUNDS, SIMPLE WORDS, TWO OR MORE WORD PHRASES) DURING THE WEEK PRIOR TO THE FIRST HOME VISIT.

My child says and/or does	My child means	Why s/he communicates *

* MY CHILD COMMUNICATES IN ORDER TO :

1. protest
2. request actions/objects
3. get attention
6. label or describe
7. answer
8. ask questions

4. imitate
5. greet

LANGUAGE COMPREHENSION CHECKLIST

PLEASE CIRCLE THE WORDS THAT YOUR CHILD CANNOT IDENTIFY OR UNDERSTAND:

OBJECTS

Food	*Household Objects*	*Outside Objects*	*Toys*
apple	tub/bath	tree	bus
milk	bed	rain	truck
juice	TV	dog	train
soup	sofa	cat	book
banana	table	plane	ball
water	chair	car	doll
cookie	room	bus	
	light		

Body parts		*Clothing*	*Important people*
	floor		
head	paper	sock	Mommy
hand	cup	shirt	Daddy
legs	dish	coat	Siblings
eyes	spoon	shoe	Baby
nose	brush	hat	child's name
hair			Grandma
foot			Gandpa
toes			name of pets

WORDS THAT DESCRIBE
hot more all gone dirty my nice did it

WORDS THAT EXPRESS FEELINGS
kiss/hug tired happy sad

SOCIAL WORDS
oh-oh hi, hello bye-bye okay nite-nite no yes

ACTION WORDS
brush blow dance cry wipe cry open kiss

stop it giddiup come peek touch run walk come
sleep push drink clap close wash sing
roll hug gimme read help pour eat

LOCATION WORDS
here there down up

QUESTIONNAIRE

1.What toys and games excite your child's interest ?

____

____

2. What books, songs and other activities capture and hold your child's attention?

____

____

3.What foods does your child prefer ?

____

____

4.Is there anything special that I should know about your child ?

____

____

APPENDIX 2

ADAPTED MIT PROTOCOL

ADAPTED MIT PROTOCOL

N.B. Symbols used are: (C) child, (R) researcher, (MM) mime or movement, (D) drum, (P) puppet, (V) visual representation of target phrase

MELODIC GROUP

Level One - Unison response

Stimulus: Once the target word or phrase has been elicited, (R) shows (C) the (V), (R) models the target twice then signals (C) to join in with (MM)

Response: (C) and (R) intone target phrase with (MM)

Progression: (C) should imitate the (MM) but may or may not imitate the words, as long as there is an attempt at singing the melodic pattern of the target phrase. Discontinue phrase if there is no attempt to respond vocally.

Level Two - Imitative response

Stimulus: (R) signals (C) to listen and watch, (R) models the target phrase with (P), supporting the rhythm of the speech pattern on (D)[13], (R) then signals (C) to repeat it. (R) cues (C) for initiation of response if necessary.

Response: (C) repeats phrase with (D)

Progression: Discontinue phrase if (C) consistently fails to produce the target after more than 4 models of the phrase.

Level Three - Conversational response

Stimulus: (R) intones question (i.e. Where's the ball ?)

Response: (C) replies with appropriate phrase with (MM)

Progression: Discontinue phrase if (C) fails after more than 3 cues to produce the desired target phrase.

---

13 The drum and hand puppet may or may not be used, depending on the play context.

ADAPTED MIT PROTOCOL

SPOKEN GROUP

The same protocol was followed as with the melodic group but without the melodic component of singing or intoning the target phrase. The (R) modelled the target phrase by <u>saying</u> it, starting at Level One.

Level One - Unison response

<u>Stimulus</u>: Once the target word or phrase has been elicited, (R) shows (C) the (V), (R) models the target twice then signals (C) to join in with (MM)

<u>Response</u>: (R) and (C) say target phrase with (MM)

<u>Progression</u>: (C) should imitate the (MM) but may or may not imitate the words as long as there is an attempt at approximating the sounds of the words. Discontinue phrase there is no attempt to approximate the sounds of the words.

Level Two - Imitative response

<u>Stimulus</u>: (R) signals (C) to listen and watch, (R) models the target phrase with (P) supporting the rhythm of the speech pattern on (D), (R) then signals (C) to repeat it. (R) cues (C) for initiation of response if necessary.

<u>Response</u>: (C) repeats phrase with (D)

<u>Progression</u>: Discontinue phrase if (C) consistently fails to produce the target after more than 4 models of the phrase.

Level Three - Conversational response

<u>Stimulus</u>: (R) asks a question (i.e. Where's the ball ?)

<u>Response</u>: (C) replies with appropriate phrase with (MM)

<u>Progression</u>: Discontinue phrase if (C) fails after more than 3 cues to elicit the desired target phrase.

APPENDIX 3

SESSION DATAFORM

SESSION DATA FORM

NAME OF CHILD: ____________________________

DATE: _______________________ SESSION: _______________

Play context	Target word	Model	Unison	Imitative	Conversation	Spontaneou

APPENDIX 4

CHAT-CODED LANGUAGE TRANSCRIPT

EXTRACT FROM A CHAT-CODED LANGUAGE TRANSCRIPT

*CHI:	alysh(a) sit dere [: there].
*ALY:	sit there?
*CHI:	yeah.
*DEB:	come close to the +/.
*CHI:	xxx.
*MOT:	come close there to alysha.
*DEB:	keep it down # ask her a question like # [/]	like take the ball and xx it say what should I do?
%add:	to ALY
*ALY:	what to do andy with the ball?
*CHI:	do this.
%com:	CHI imitates the action of throwing the ball
*ALY:	what to do with the ball what do I do with it?
*CHI:	roll it!
*ALY:	roll it?
*CHI:	yeah.
*CHI:	like dat.
*CHI:	no.
*CHI:	do this.
*ALY:	roll straight [?].
*DEB:	yeah say what else +/.
*ALY:	andy you throw it # hmm?
*CHI:	in guy [: sky]!
*ALY:	in the sky?
*CHI:	yeah!
*MOT:	you ask alysha shall I throw it or pass it # <or	roll it> [>] ?
*CHI:	xxx [<].
*CHI:	<pass it> [>].
*MOT:	<no you ask her> [<].
*ALY:	pass it.
*MOT:	no you ask alysha.
*CHI:	pass [?] alysha.
*CHI:	yeah.
%com:	all laugh and whisper
*ALY:	roll it pass it?
*CHI:	xx.
*ALY:	in the sky okay.

*DEB: okay # what should I do # hmm who should I throw it to?
*CHI: (a)lysha.

REFERENCES

Anonymous (1994). Assessment: Melodic intonation therapy. Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology. *Neurology, 44,* 566-568.

Ayres, A.J.(1979). *Sensory integration and the child.* Los Angeles, CA: Western Psychological Services.

Berlin, C.I. (1976). On: Melodic intonation therapy for aphasia. *Journal of Speech and Hearing Disorders, 41*, 298-299.

Blau, A., Lahey, M. & Oleksiuk-Velez, A. (1984). Planning goals for intervention: Testing or language sampling ? *Exceptional Children, 51*(1), 78-79.

Bloom, L. & Lahey, M. (1978). *Language development and language disorders*. N.Y: John Wiley.

Bogdan, R.C. and Biklen, S.K. (1991). *Qualitative research for education: An introduction to theory and methods.* Newton, Mass: Allyn & Bacon.

Bordens, K. & Abbott, B. (1991). *Research designs and methods: A process approach* (2nd ed.). Mountain View, CA: Mayfair.

Bricker, D.D. & Carlson, R. (1981). Issues in Early Language Intervention. In R.L. Shiefelbusch and D.D. Bricker (Eds.). *Early language: Acquisition and intervention, Language intervention series, Volume VI.* Baltimore, Maryland: University Park Press.

Brown, R. (1973). *A first language: The early years*. Cambridge MA: Harvard University Press.

Bryant, D.M.& Graham, M.A. (Eds).(1993). *Implementing early intervention.* N.Y: Guilford Press.

Campbell, D.T. & Stanley, J.C. (1979). *Experimental and quasi-experimental designs for research* Boston: Houghton Mifflin.

Carroll, D. (1989). *Music therapy and the language-delayed child.* Proceedings of the 16th Annual Conference of the Canadian Association for Music Therapy.

Carroll, D. (1992). *Designing a research study in music therapy - From field observations to research proposal.* Proceedings of the 19th Annual Conference of the Canadian Association for Music Therapy.

Cartwright, J. & Huckaby, G. (1972). Intensive preschool language program. *Journal of Music Therapy, 9,* 137-146.

Cicchetti, D. & Beeghly, M. (Eds.). (1990).*Children with Down syndrome - A developmental perspective.* Cambridge: University Press.

Clunies-Ross, G.G.(1979). Accelerating the development of Down syndrome infants and young children. *Journal of Special Education, 13(2),* 169-177.

Clynes, M. (Ed.). (1982). *Music, mind and brain: The neuropsychology of music.* New York: Plenum Press.

Cohen, N. & Masse, R. (1993). The application of singing instruction on the speech production of neurologically impaired persons. *Journal of Music Therapy, 30,* 81-99.

Cohen, N. (1994). Speech and song: Implications for therapy, *Music Therapy Perspectives, 12(1),* 8-14.

Conti-Ramsden G. and Snow, C.E. (1990). Children's language: How it develops and how it is used. In Conti- Ramsden G. and Snow, C.E. (Eds.). *Children's Language Volume 7.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Cornwell, A. & Birch, H. (1969). Psychological and social development in home-reared children with Down Syndrome *Americal Journal of Mental Deficiency,74,* 341-350.

Damasio A.R. & Damasio, H. (1977). Musical faculty and cerebral dominance. In M. Critchley & R.A. Henson (Eds.). *Music and the brain: Studies in the neurology of music.* Southhampton: Camelot Press.

Darley, F.L. (1982). .*Aphasia*. Philadelphia: W.B.Saunders, Co.

Darrow, A. (1990). Beyond mainstreaming - Dealing with diversity. *Music Educators Journal, 78,* 4, 36-39.

Ditson, R.(1961). A study of the effects of moderate background music on the behaviour of cerebral palsied children. *Bulletin of the National Association for Music Therapy, 10* (6).

Duchan, J. & Lund, N.(1988). *Assessing children's language in naturalistic contexts*. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Fernald, A. (1989). Intonation and communicative intent in mother's speech to infants: Is the melody the message ? *Child Development, 60,* 1497-1510.

Fey, M. (1986). *Language intervention with young children*. Boston, Mass.: College-Hill Publication.

Florey, L.(1971). An approach to play and play development. *The American Journal of Occupational Therapy, Vol.25* (6), 275-280.

Fowler, A.E. (1990). Language abilities in children with Down syndrome: Evidence for a specific syntactic delay. In D. Cicchetti & M. Beeghly (Eds.), *Children with Down syndrome - A developmental perspective.* Cambridge: University Press.

Gates, A. & Bradshaw, J.L. (1977). The role of the cerebral hemispheres in music. *Brain and Language, 4,* 403-431.

Goodglass, H. & Calderon, M. (1977). Parallel processing of verbal and musical stimuli in right and left hemisphere. *Neuropsychologia, 15*, 397-407.

Gottsleben & Tyack (1974). *Language sampling, analysis & training.*

Graham, R.(1988). Barrier-free music education: Methods to make mainstreaming work. *Music Educators Journal, 76,* 1, 29-33.

Harding, C. & Ballard, K.(1982). The effectiveness of music as a stimulus and as a contingent reward in promoting the spontaneous speech of three physically handicapped preschoolers. *Journal of Music Therapy, 19,* 86-101.

Helfrich-Miller, K. (1980). *The use of melodic intonation therapy with developmentally apraxic children.* Paper presented at the convention of the American Speech and Hearing Association, Detroit, Michigan, November, 1980.

Hodapp, R.M. & Zigler, E. (1990). Applying the developmental perspective to individuals with Down syndrome. In D. Cicchetti & M. Beeghly (Eds.), *Children with Down syndrome - A developmental perspective.* Cambridge: University Press.

Hoskins, C. (1988). The use of music to increase verbal response and to improve expressive language abilities in preschool language-delayed children. *Journal of Music Therapy, 25,* 73-88.

Hoshizaki, M.K. (1983). *Teaching mentally retarded children through music.* Springfield, Mass. C.C. Thomas Publishers.

Humpal, M.E. (1990). Early intervention: The implications for music therapy. *Music Therapy Perspectives, 8,* 30-35.

Irwin, C.E. (1971). *The use of music in a speech and language development program with mentally* *retarded Down's syndrome children.* Unpublished thesis. Tallahassee, Florida: Florida State University.

Jellison, J.A. & Gainer, E.W. (1995). Into the mainstream: A case study of a child's participation in music education and music therapy. *Journal of Music Therapy, 22*(4), 228-247.

Kazdin, A.E. (1980). *Research design in clinical psychology.* New York: Harper & Row Publishers.

Krauss, T. & Galloway, H. (1982). Melodic intonation therapy with language-delayed apraxic children. *Journal of Music Therapy, 19,* 102-113.

Lathom, W., Edson, S. & Toombs, M.R. (1965). A coordinated speech therapy and music therapy program. *Journal of Music Therapy, 2,* 188-120.

Leung, K.(1985). *Enhancing the speech and language development of communicatively disordered children through music and movement.* Paper presented at the annual convention of the Council for Exceptional Children, Anaheim, CA., April, 1985.

Lott, T.L. & McCoy, E.E. (Eds.)(1992). *Down syndrome: Advances in medical care.* New York: Wiley-Liss, Inc.

Lunt, I. (1973). Rhythm and the slow learner, *Special Education, 62*, 21-23.

Luria, A.R. (1966). *Speech and the development of mental processes in the child.* London, England: Staples Press.

MacWhinney, B. (1991). The *CHILDES project: Tools for analysing talk.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Mahoney, G. & Snow, K. (1983). The relationship of sensorimotor functioning to children's response to early language training. *Mental Retardation,* 21, 248-254.

Manolson, A. (1992). *It takes two to talk: A parent's guide to helping children communicate.* Toronto, Ont.: A Hanon Centre Publication

Marshall, N. & Holtzapple, P. (1976). Melodic intonation therapy: Variations on a theme. In R.H. Brookshire (Ed.), *Clinical Aphasiology, Proceedings of the Conference,* Minneapolis, Minn.: BRK Publishers.

Martin, J.G. (1972). Rhythmic (Hierarchical) versus serial structure in speech and other behavior. *Psychological Review, 79(6),* 487-509.

McLean, J.E. & Snyder-McLean, L.K.(1978). *A transactional approach to early language stimulation.* Merril, Communication Development and Disorders Series.

Merriam, S.B. (1990). *Case study research in education: A qualitative approach*. San Francisco, Calif.: Jossy-Bass Publishers.

Michel, D. (1974). Development of music therapy procedures with speech language disorders. *Journal of Music Therapy, 14,* 74-80.

Miles, M.B. & Huberman, A.M. (1993). *Qualitative data analysis: An expanded sourcebook.* Newbury Park, Calif.: Sage Publications.

Miller, J.F.(1992). Development of speech and language in children with Down syndrome. In T.L. Lott & E.E. McCoy Eds.). *Down syndrome: Advances in medical care*, New York: Wiley-Liss, Inc.

Miller, S.B. & Toca, J.M. (1979). Adapted melodic intonation therapy: A case study of an experimental language program for an autistic child. *Journal of Clinical Psychiatry,* April, 1979, 80-82.

Nadel, L. (1992). Learning and cognition in Down syndrome. In T.L. Lott & E.E. McCoy (Eds.). *Down syndrome: Advances in medical care,* New York: Wiley-Liss, Inc.

Naeser, M.A. & Helm-Estabrooks, N. (1985). CT scan lesion localization and response to melodic intonation therapy with nonfluent aphasia cases. *Cortex, 21(2),* 203-223.

O'Boyle, M.W. & Sanford M. (1988). Hemispheric asymmetry in the matching of melodies to rhythm sequences tapped in the right and left palms. *Cortex, 24,* 211-221.

Piaget, J. (1959). *The language and thought of the child.* London: Routledge & Kegan Paul.

Piaget, J. (1969). *Judgment and reasoning in the child.* London: Routledge & Kegan Paul.

Preuss, J.B., Vadasy, P.F. & Fewell, R.R.(1987). Language development in children with Down syndrome: An overview of recent research. *Education and Training in Mental Retardation, 22(1),* 44-55.

Rogers, A., & Fleming, P.L. (1981). Rhythm and melody in speech therapy for the neurologically impaired. *Music Therapy, 1,* 33-38.

Romski, M.A.(1980). *Melodic intonation therapy: Application to a childhood motor sequencing impairment.* Paper presented at the convention of the American Speech and Hearing Association, Detroit, Michigan.

Ross, E.D. & Mesulam, M. (1979). Dominant language functions of the right hemisphere? Prosody and emotional gesturing. *Archives in Neurology, 36,* 144-148.

Scarpa, E.M. (1990). Intonation and dialogue processes in early speech. In G. Conti-Ramsden and C.E. Snow (Eds.). *Children's Language Volume 7.* Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Seybold, C.D. (1971). The value and use of music activities in the treatment of speech-delayed children. *Journal of Music Therapy, 8,* 102-110.

Sparks, R.W., Helm, N. & Albert, M.(1973). Melodic intonation therapy for aphasia. *Archives of Neurology, 29,* 130-131.

Sparks, R.W., Helm, N. & Albert, M.(1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex, 10*, 303-316.

Sparks, R.W. & Holland, A.L.(1976). Method: Melodic intonation therapy for aphasia. *Journal of Speech and Hearing Disorders, 41,* 287-297.

Sparks, R.W. & Deck, J.W. (1994). Melodic intonation therapy. In R. Chapey (Ed.), *Language intervention strategies in adult aphasia,* Baltimore, MA.: Williams & Wilkins.

Spiker, D.(1990). Early intervention from a developmental perspective. In D. Cicchetti & M. Beeghly (Eds.). *Children with Down syndrome - A developmental perspective.* Cambridge: University Press.

Swift, E. & Rosin, P. (1990). A remediation sequence to improve speech intelligibility for students with Down syndrome. *Language, Speech, and Hearing Services in Schools, 21*,140-146.

Thompson, K.P. (1990). Working toward solutions in mainstreaming. *Music Educators Journal, 78, 4,* 30-35.

Thurman, S.K. & Widerstrom, A.H. (1990). *Infants and young children with special needs- A developmental and ecological approach (*2nd ed.). Baltimore, MA: Paul H. Brookes Publishing Company, Inc.

Tudge, J.(1990). Vygotsky, the zone of proximal development, and peer collaboration: Implications for classroom practice. In Moll, L.(Ed). *Vygotsky and education: Instructional implications and applications of sociohistorical psychology, (*pp.155-172). Mass.: Cambridge University Press.

Uzgiris, I.C. (1981). Experience in the social context: Imitation and play. In R.L. Shiefelbusch and D.D. Bricker (Eds.). *Early Language: Acquisition and intervention, Language Intervention Series Volume VI.* Baltimore, Maryland: University Park Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Vygotsky, L.S. (1986). *Thought and language.* Cambridge, MA: The MIT Press.

Walker, J. (1972). The use of music as an aid in developing functional speech in the institutionalized mentally retarded. *Journal of Music Therapy, 9,* 1-12.

Warwick, A. (1988). Questions and reflections on research. *Journal of British Music Therapy, 2*(2), 7-11.

Winzer, M. (1990). *Children with exceptionalities - A canadian perspective* (2nd ed.). Scarborough, Ont.: Prentice- Hall Canada, Inc.

Zoller, M.B. (1991). Use of music activities in speech-language therapy. *Language, Speech and Hearing Services in Schools, 22,* 272-276.

From Singing to Speaking: Why Singing May Lead to Recovery of Expressive Language Function in Patients with Broca's Aphasia

Gottfried Schlaug, Sarah Marchina, & Andrea Norton
Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts

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

Of 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;

Music Perception volume 25, issue 4, pp. 315–323, issn 0730-7829, electronic issn 1533-8312 DOI:10.1525/mp.2008.25.4.315

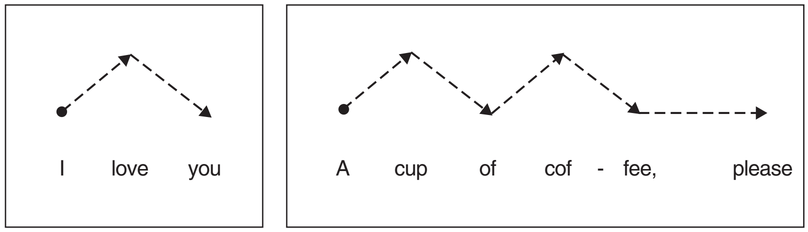


FIGURE 1. Spoken phrases (prosodic patterns) transposed into melodic intonation patterns. Pitches are determined by the natural prosody of speech—accented syllables are presented on the higher of the two pitches.

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 aphasia 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 fronto-temporal 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) aphasic patients (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 therapist 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 tapping that may likely be the key factors in its effectiveness.

The original interpretation of MIT's path to successful 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 hemispheres. Belin et al. (1996) suggested that MIT-facilitated recovery was associated with the reactivation of left-hemisphere regions, most notably the left prefrontal cortex, just anterior to Broca's region. Although, this publication was the first to examine patients treated with an MIT-like intervention using functional neuroimaging, 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 colleagues' findings it is important to consider the following: 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 diagnosed 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 superior 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 traditional 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 living, 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; 1-2 years of instrumental practice as a child and some singing in choirs in high school and college; moderate to severe right hemiparesis; independent in activities of daily living, ADL; Barthel-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 confrontational 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, Zaehle, 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).

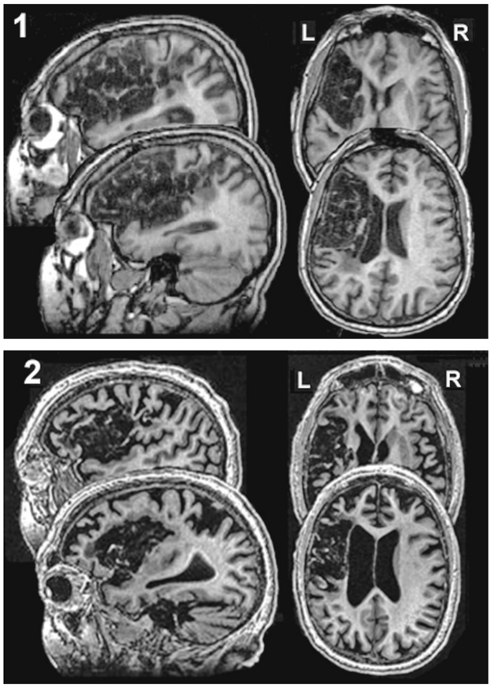


FIGURE 2. High-resolution, T1-weighted images show the chronic, left-hemisphere lesion location and extent of Patients #1 and #2, encompassing both Broca's region and the anterior part of the superior temporal lobe.

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.

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

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 Table1).

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, & Gangittano, 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

Albert, M. L., Sparks, R. W., & Helm, N. A. (1973). Melodic intonation therapy for aphasia. *Archives of Neurology, 29,* 130-131.

Aziz-Zadeh, L., Wilson, S. M., Rizzolatti, G., & Iacoboni, M. (2006). Congruent embodied representations for visually presented actions and linguistic phrases describing actions. *Current Biology, 16,* 1818-1823.

Basso, A., Gardelli, M., Grassi, M. P., & Mairotti, M. (1989). The role of the right hemisphere in recovery from aphasia. Two case studies. *Cortex 25,* 555-566.

Belin, P., Van Eeckhout, P., Zilbovicius, M., Remy, P., François, C., Guillaume, S., et al. (1996). Recovery from nonfluent aphasia after melodic intonation therapy: A PET study. *Neurology, 47,* 1504-1511.

Blasi, V., Young, A. C., Tansy, A. P., Petersen, S. E., Snyder, A. Z., & Corbetta, M. (2002). Word retrieval learning modulates right frontal cortex in patients with left frontal damage. *Neuron, 36,* 159-170.

Bohland, J. W., & Guenther, F. H. (2006). An fMRI investigation of syllable sequence production. *Neuroimage, 32,* 821-841.

Bonakdarpour, B., Eftekharzadeh, A., & Ashayeri, H. (2000). Preliminary report on the effects of melodic intonation therapy in the rehabilitation of Persian aphasic patients. *Iranian Journal of Medical Sciences, 25,* 156-160.

Boucher, V., Garcia, L. J., Fleurant, J., & Paradis, J. (2001). Variable efficacy of rhythm and tone in melody-based interventions: implications for the assumption of a right-hemisphere facilitation in nonfluent aphasia. *Aphasiology, 15,* 131-149.

Broca, P. (1864) Deux cas d'aphemie tramatique, produite par des lesions de la troisieme circonvolution frontale gauche. *Bulletin de la Societe de Chirurgie de Paris, V,* 51-54.

Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., & Parsons, L. M. (2004). The song system of the human brain. *Brain Research Cognitive Brain Research, 20,* 363-75.

Cappa, S. F., Perani, D., Grassi, F., Bressi, S., Alberoni, M., Franceschi, M., et al. (1997). A PET follow-up study of recovery after stroke in acute aphasics. *Brain and Language, 56,* 55-67.

Cappa, S., & Vallar, G. (1992). The role of left and right hemispheres in recovery from aphasia. *Aphasiology, 6,* 356-372.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic.

Cohen, N. S., & Masse, R. (1993). The application of singing and rhythmic instruction as a therapeutic intervention for persons with neurogenic communication disorders. *Journal of Music Therapy, 30,* 81-99.

Cornelissen, K., Laine, M., Tarkiainen, A., Järvensivu, T., Martin, N., & Salmelin, R. (2003). Adult brain plasticity elicited by anomia treatment. *Journal of Cognitive Neuroscience, 15,* 444-461.

de Groot, A. D. (1965). *Thought and choice in chess*. The Hague, The Netherlands: Mouton & Co.

Gaab, N., Gaser, C., Zaehle, T., Chen, Y., & Schlaug, G. (2003). Functional anatomy of pitch memory—A fMRI study with sparse temporal sampling. *Neuroimage, 19,* 1417-1426.

Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Language and motor control. *Experimental Brain Research, 133,* 468-490.

Gerstman, H. L. (1964). A case of aphasia. *Journal of Speech & Hearing Disorders, 29,* 89-91.

Geschwind, N. (1971). Current concepts: Aphasia. *New England Journal of Medicine, 284,* 654-656.

Goodglass, H., & Kaplan, E. (1983). *Boston Diagnostic Aphasia Examination* (2nd ed.). Philadelphia: Lea & Febiger.

Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review 105,* 611-633.

Hebert, S., Racette, A., Gagnon, L., & Peretz, I. (2003). Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain, 126,* 1-13.

Heiss, W. D., Kessler, J., Thiel, A., Ghaemi, M., & Karbe, H. (1999). Differential capacity of left and right hemispheric areas for compensation of poststroke aphasia. *Annals of Neurology,45,* 430-438.

Heiss, W. D., & Thiel, A. (2006). A proposed regional hierarchy in recovery of post-stroke aphasia. *Brain and Language, 98,* 118-123.

Helm-Estabrooks, N., & Albert, M. L. (1991). *Manual of aphasia therapy.* Austin: Pro-Ed.

Helm-Estabrooks, N., Nicholas, M., & Morgan, A. (1989). *Melodic intonation therapy.* Austin: Pro-Ed.

Jeffries, K. J., Fritz, J. B., & Braun, A. R. (2003). Words in melody: An $H_2 15O$ PET study of brain activation during singing and speaking. *Neuroreport, 14,* 749-745.

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test* (2nd ed.). Baltimore and Philadelphia: Lippincott, Williams & Wilkins.

Keith, R. L., & Aronson, A. E. (1975). Singing as therapy for apraxia of speech and aphasia: Report of a case. *Brain and Language, 2,* 483-488.

Kertesz, A., Lesk, D., & McCabe, P. (1977). Isotope localization of infarcts in aphasia. *Archives of Neurology, 100,* 1-18.

Kinsbourne, M. (1998). The right hemisphere and recovery from aphasia. In B. Stemmer & H. A. Whitaker (Eds.), *Handbook of neurolinguistics* (pp. 386-393). New York: Academic Press.

Kinsella, G., Prior, M. R., & Murray, G. (1988). Singing ability after right and left sided brain damage. A research note. *Cortex, 24,* 165-169.

Koelsch, S., Gunter, T. C., von Cramon, D. Y., Zysset, S., Lohmann, G., & Friederici, A. D. (2002). Bach speaks: A cortical "language-network" serves the processing of music. *Neuroimage, 17,* 956-966.

Koelsch, S., Fritz, T., Schulze, K., Alsop, D., & Schlaug, G. (2005). Adults and children processing music: An fMRI study. *Neuroimage, 25,* 1068-1076.

Lahav, A., Saltzman, E., & Schlaug, G. (2007). Action representation of sound: Audiomotor recognition network while listening to newly-acquired actions. *Journal of Neuroscience, 27,* 308-314.

Meister, I. G., Boroojerdi, B., Foltys, H., Sparing, R., Huber, W., & Topper, R. (2003). Motor cortex hand area and speech: Implications for the development of language. *Neuropsychologia, 41,* 401-406.

Mimura, M., Kato, M., Sano, Y., Kojima, T., Naeser, M., & Kashima, H. (1998). Prospective and retrospective studies or recovery in aphasia—Changes in cerebral blood flow and language functions. *Brain, 121,* 2083-2094.

Moore, W. H. (1989). Language recovery in aphasia: A right hemisphere perspective. *Aphasiology, 3,* 101-110.

Musso, M., Weiller, C., Kiebel, S., Muller, S. P., Bulau, P., & Rijntjes, M. (1999). Training-induced brain plasticity in aphasia. *Brain, 122,* 1781-1790.

Ozdemir, E., Norton, A., & Schlaug, G. (2006). Shared and distinct neural correlates of singing and speaking. *Neuroimage, 33,* 628-635.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6,* 674-681.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain, 113,* 1185-205.

Rosen, H. J., Petersen, S. E., Linenweber, M. R., Snyder, A. Z., White, D. A., Chapman, L., et al. (2000). Neural correlates of recovery from aphasia after damage to left inferior frontal cortex. *Neurology, 55,* 1883-1894.

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain, 129,* 1371-1384.

Schuppert, M., Munte, T. F., Wieringa, B. M., & Altenmuller, E. (2000). Receptive amusia: Evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain, 123,* 546-559.

Selnes, O. A. (1999). Recovery from aphasia: Activating the "right" hemisphere. *Annals of Neurology, 45,* 419-420.

Small, S. L., Flores, D. K., & Noll, D. C. (1998). Different neural circuits subserve reading before and after therapy for acquired dyslexia. *Brain and Language, 62,* 298-308.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174-215.

Sparks, R., Helm, N., & Albert, M. (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex, 10,* 303-316.

Sparks, R., & Holland, A. (1976). Method: Melodic intonation therapy for aphasia. *Journal of Speech and Hearing Disorders 41,* 287-297.

Thaut, M. H., Kenyon, G. P., Schauer, M. L., & McIntosh, G. C. (1999). Rhythmicity and brain function: Implications for therapy of movement disorders. *IEEE Engineering in Medicine and Biology Magazine, 18,* 101-108.

Thompson, C. K., & Shapiro, L. P. (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms. *Aphasiology 19,* 10-11.

Thompson, C. K., & Shapiro, L. P. (2007). Complexity in treatment of syntactic deficits. *American Journal of Speech and Language Pathology, 16,* 30-42.

Tokimura, H., Tokimura, Y., Oliviero, A., Asakura, T., & Rothwell, J. C. (1996). Speech-induced changes in corticospinal excitability. *Annals of Neurology, 40,* 628-634.

Uozumi, T., Tamagawa, A., Hashimoto, T., & Tsuji, S. (2004). Motor hand representation in cortical area 44. *Neurology, 62,* 757-761.

Warburton, E., Price, C. J., Swinburn, K., & Wise, R. J. (1999). Mechanisms of recovery from aphasia: evidence from positron emission tomography studies. *Journal of Neurology, Neurosurgery and Psychiatry, 66,* 155-161.

Weiller, C., Isensee, C., Rijntjes, M., Huber, W., Müller, S., Bier, D., et al. (1995). Recovery from Wernicke's aphasia: A positron emission tomographic study. *Annals of Neurology, 37,* 723-732.

Wilson, S. J., Parsons, K., & Reutens, D. C. (2006). Preserved singing in aphasia: A case study of the efficacy of the Melodic Intonation Therapy. *Music Perception, 24,* 23-36.

Winhuisen, L., Thiel, A., Schumacher, B., Kessler, J., Rudolf, J., Haupt, W. F., & Heiss, W. D. (2005). Role of the contralateral inferior frontal gyrus in recovery of language function in poststroke aphasia: A combined repetitive transcranial magnetic stimulation and positron emission tomography study. *Stroke, 36,* 1759-1763.

Zatorre, R. J., & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex, 11,* 946-953.

Preserved Singing in Aphasia: A Case Study of the Efficacy of Melodic Intonation Therapy

Sarah J. Wilson
School of Behavioural Science, University of Melbourne, & Department of Clinical Neuropsychology, Austin Health

Kate Parsons
School of Behavioural Science, University of Melbourne

David C. Reutens
Department of Neurology, Southern Health Neurosciences, Monash Medical Centre

This study examined the efficacy of Melodic Intonation Therapy (MIT) in a male singer (KL) with severe Broca's aphasia. Thirty novel phrases were allocated to one of three experimental conditions: unrehearsed, rehearsed verbal production (repetition), and rehearsed verbal production with melody (MIT). The results showed superior production of MIT phrases during therapy. Comparison of performance at baseline, 1 week, and 5 weeks after therapy revealed an initial beneficial effect of both types of rehearsal; however, MIT was more durable, facilitating longer-term phrase production. Our findings suggest that MIT facilitated KL's speech praxis, and that combining melody and speech through rehearsal promoted separate storage and/or access to the phrase representation.

Received August 26, 2005, accepted: March 14, 2006

Key words: aphasia, amusia, music therapy, melodic intonation therapy, nonpropositional and propositional speech

The last 250 years have seen a number of reports describing preserved singing (with lyrics) in aphasia (Brust, 1980, 2001). This phenomenon raises important questions about the representation of music and language functions in the brain. For example, Yamadori and colleagues (1977) found that out of 24 patients with Broca's aphasia, 12 were able to sing a well-known song with text correctly. The absence of a relationship between the severity of aphasia and text production during singing led them to conclude that singing and speech are independent functions. Considerable previous research now suggests that singing is predominantly a right-hemisphere function while speech is primarily mediated by the left hemisphere (Botez & Wertheim, 1959; Epstein et al., 1999; Geschwind, Quadfasel, & Segarra, 1968; Gleason & Goodglass, 1984; Gordon & Bogen, 1974; Jackson, 1871; Perry, Zatorre, Petrides, Alivisatos, Meyer, & Evans, 1999; Riecker, Ackermann, Wildgruber, Dogil, & Grodd, 2000; Smith, 1966; Stewart, Walsh, Frith, & Rothwell, 2001; Wildgruber, Ackermann, Klose, Kardatzki, & Grodd, 1996).

One explanation for preserved singing in aphasia is that song is a form of nonpropositional speech (Ellis & Young, 1996). First described by Jackson (1866), the distinction between propositional and nonpropositional speech is commonly used in studies of aphasic language. Jackson (1915) defined nonpropositional speech as ready-made utterances that may express emotion but are "intellectually dead," while propositional speech is an intentional expression of thought. He proposed that in the event of left hemisphere damage affecting propositional speech, sensorimotor processes in the right hemisphere may still allow nonpropositional speech (Jackson, 1874). Recent researchers have built on Jackson's ideas, defining nonpropositional speech as conventionalized, context-dependent speech that does not involve syntactic parsing or the conscious formulation of new utterances to express a semantic message (Code, 1997; Nenonen, Niemi, & Laine, 2002; Speedie, Wertman, Ta'ir, & Heilman, 1993). The notion that singing represents a form of nonpropositional speech has received variable support (Hébert, Racette, Gagnon & Peretz, 2003). Notably, Speedie and colleagues (1993) described a case with preserved speech but impaired ability to swear, recite poetry, and sing previously well-known songs following a right basal ganglia lesion.

Music Perception volume 24, issue 1, pp. 23–36, issn 0730-7829, electronic issn 1533-8312

Proponents of Melodic Intonation Therapy (MIT) claim that pairing melody with words facilitates propositional speech in aphasic patients (Naeser & Helm-Estebrooks, 1985). A number of explanations have been proposed for this effect, the original being that melody activates right hemisphere language structures that were formerly dominated by the left hemisphere (Albert, Sparks, & Helm, 1973). A related account is that MIT increases the role of the right hemisphere in interhemispheric language control (Sparks, Helm, & Albert, 1974). Neither of these mechanisms has received recent support, however, in a positron emission tomography study of recovery from nonfluent aphasia after MIT. Specifically, Belin and colleagues (1996) found that the repetition of words with MIT reactivated Broca's area and the left prefrontal cortex while deactivating the homologue of Wernicke's area in the right hemisphere.

Donnan and colleagues (1999) have proposed that Broca's aphasia involves three different levels of language disorder: initial mutism, speech dyspraxia, and agrammatism. An alternative account of the mechanism of MIT is speech facilitation at the articulatory level, through the use of a slower tempo, a more precise rhythm, and greater accentuation than normal speech (Sparks & Holland, 1976). Slowing the tempo can lessen the neuromotor complexity of the articulatory process, as evidenced by the beneficial effect of melody on stuttering (Ham, 1990). Furthermore, syllable lengthening has been shown to aid phrase production in nonfluent aphasics using MIT (Laughlin, Naeser, & Gordon, 1979).

Despite a number of papers advocating the use of MIT (Albert et al., 1973; Baker, 2000; Naeser & Helm-Estebrooks, 1985) very few studies have employed experimental designs that allow the various mechanisms of MIT to be compared. More fundamental to this, very few studies have compared MIT with control tasks or accounted for time since injury, making it difficult to distinguish spontaneous recovery from improvements specifically related to MIT. Methodological weaknesses in the literature such as these have recently led Hébert and colleagues (2003) to question the empirical basis of the proposed beneficial effects of MIT.

Three previous studies employing control tasks have failed to show a facilitatory effect of melody on speech production. Specifically, Cohen and Ford (1995) found no differences in the speech production of 12 aphasic patients compared across three experimental conditions; verbal production only, verbal production with rhythm, and verbal production with melody. More recently, Hébert and colleagues (2003), and Peretz, Gagnon, Hébert, and Macoir (2004) reported no differences between the mean number of sung or spoken words produced by two nonfluent aphasics when using either familiar or novel stimuli. In all of these studies, however, MIT was not administered as part of the experimental design, preventing direct assessment of the efficacy of MIT.

The aim of the present study was to evaluate the efficacy of MIT using a pre- versus post-treatment design that included control tasks in a neurologically *stable* male singer (KL) with severe Broca's aphasia. KL is a right-handed amateur musician who retained the ability to sing familiar songs with lyrics following a large left frontoparietal infarct. His neurologically preserved singing provided the ideal opportunity to assess the ability of MIT to promote the production of propositional speech. Thirty novel phrases were generated and allocated to one of three experimental conditions: unrehearsed, rehearsed verbal production (repetition), and rehearsed verbal production with melody (MIT). The unrehearsed condition served as the control for the rehearsed conditions (i.e., the effect of no intervention). The rehearsed conditions entailed twice-weekly practice sessions for a period of 4 weeks under the guidance of a Registered Music Therapist. Rehearsed verbal production assessed the effects of practice using an accentuated rhythm as opposed to melody during training. The immediate and longer-term efficacy of MIT was assessed by comparing KL's performance at baseline, 1 week, and 5 weeks after therapy. The results are discussed in light of the proposed mechanisms of MIT.

Methods

Case KL

KL was a right-handed amateur male musician born in 1949. Following informed consent from KL and his sister, he participated in this study between March 2001 and June 2002. In May 1997, KL sustained a left middle cerebral artery tertiary stroke. Initial examination revealed a dense right hemiparesis and global aphasia. Since 1998, he has suffered infrequent focal motor seizures affecting the right hand with infrequent secondarily generalization.

Magnetic resonance imaging (MRI) performed 3 years post-injury revealed dilatation of the left lateral ventricle and an infarct extending anteriorly to the middle frontal gyrus and posteriorly to the inferior parietal lobe. Marked cortical thinning of the left middle and superior temporal gyri and insular cortex and damage to the left basal ganglia were evident (see Figure 1).

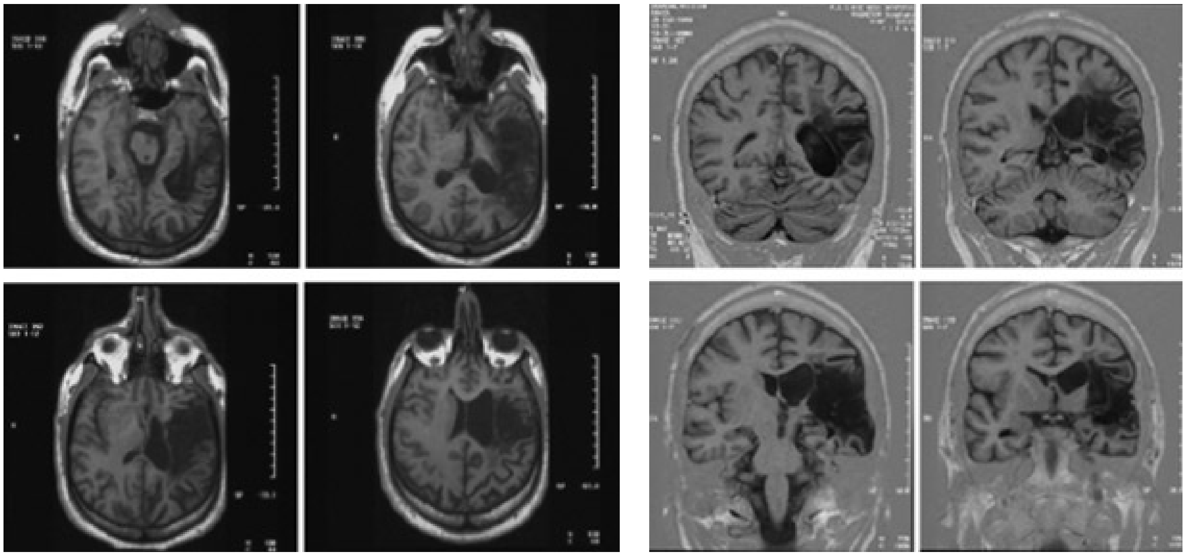


FIG. 1. Axial and coronal magnetic resonance imaging (MRI) scans of KL's left frontoparietal lesion, taken 3 years post-stroke. The right hemisphere appears on the left-hand side of the scan.

Four years post-injury, KL presented with a stable, severe expressive aphasia. His spontaneous speech was limited to counting to 10 and simple words that were produced with preserved articulation and prosody (e.g., "yes," "no," "look"). At a conversational level, KL showed preserved language comprehension and was able to communicate through gesture, facial expression and the use of props (e.g., photographs). He required the listener to establish the topic, ask yes/no questions, and clarify the message given or received.

KL was a self-taught musician and was the lead singer of a blues/rock band for 25 years prior to his stroke. He also played the piano and the harmonica, and had composed approximately 40 original songs for the band. Post-stroke he retained the ability to hum, and 1 year post-injury he was able to sing familiar songs with lyrics following practice with a music therapist and friends.

NEUROPSYCHOLOGICAL PROFILE

KL was university educated and taught English and Politics at secondary school for 30 years prior to his stroke. He had a performance IQ (PIQ) of 79 as measured on the Wechsler Adult Intelligence Scale-III (Wechsler, 1997). This was below his expected premorbid level but consistent with the location and extent of his infarct. In particular, KL's low PIQ reflected marked difficulties with psychomotor speed (Processing Speed Index = 69) and more subtle difficulties with visuospatial functioning (Perceptual Organization Index = 86). In contrast, his nonverbal memory function measured using the Wechsler Memory Scale-Revised (Wechsler, 1987) fell within the normal range (Visual Memory Index = 95, Visual memory span of six items forward and backward).

The Boston Diagnostic Aphasic Examination (BDAE; Goodglass & Kaplan, 1972) revealed that KL's language problems are primarily related to expression; however, he also showed comprehension difficulties typical of classical Broca's aphasia (see Table 1). In the expressive domain, KL showed poor verbal agility. His articulation was normal only for a limited number of familiar words that he could use in phrases no longer than two words (e.g., "Yes, but . . ."). He displayed evidence of verbal apraxia, correctly repeating words only by closely copying the speaker's articulation. He was unable to repeat or read aloud phrases, name objects, or respond to simple questions even though he demonstrated he knew the answers. His writing skills paralleled his spoken speech; he could copy words but not compose his own sentences. He was able to write his name and serial sequences without copying, but could not write to dictation or perform the more complex written tasks of the BDAE.

In the receptive domain, KL's comprehension difficulties were most evident on the Ideational material task, which requires a yes/no response to questions unrelated to the immediate context (e.g., "Will a stone sink in water?"). In contrast, he could follow group conversations quite easily, highlighting his use of contextual cues to understand speech. He showed a strength on the performance of written comprehension tasks, correctly matching printed words with spoken words or pictures, and choosing the appropriate word to 'fill the

TABLE 1. Assessment of KL's language skills using the Boston Diagnostic Aphasic Examination

Task	Score	Comments
Oral expression		
Verbal agility[a]	3/14	
Repeating words	8/10	
Repeating phrases	1/16	
Reading aloud words[b]	4/10	
Reading aloud phrases[b]	0/10	
Answering simple questions	0/10	e.g., "What do we tell the time with?"
Naming	0/35	
Body-part naming	0/10	
Written expression		
Name and address	✓	
Serial sequences		
Alphabet	15/26	
Numbers 1–21	21/21	
Dictation	2/15	Included letters, numbers and words
Auditory comprehension		
Body part discrimination	6/26	Difficulty with "right" and "left"
Commands	7/15	Difficulty with commands > 3 sequences
Word-picture matching		
Colors	0/6	Accurate performance of Ishihara's color vision test
Numerals	3/6	
Letters	2/6	
Objects	6/6	
Shapes	6/6	
Verbs	6/6	
Ideational material	5/12	Used contextual cues to follow informal group conversations
Written comprehension		
Word-picture matching		
Colors	2/2	
Numerals	1/2	
Shapes	2/2	
Objects	2/2	
Verbs	2/2	
Spoken-written word matching	5/8	
"Fill in the gaps"	6/10	

[a] Verbal agility is a measure of the number of times an individual can repeat words as accurately as possible in 5 seconds.
[b] KL retained the ability to read written text. Difficulties reading aloud primarily reflected his impaired speech production.

gap' in sentences. Interestingly, he showed a dissociation between the correct matching of colors to their written names but not their spoken names (see Table 1).

Experimental tasks similar to those designed by Linebarger and colleagues (1983) and Carramazza and Hillis (1989) were used to assess KL's grammatical skills. In the first task, pairs of sentences were presented both visually (printed on cards) and aurally (read aloud), and KL was required to identify the grammatically correct sentence within each pair (Linebarger, Schwartz, & Saffran, 1983). In the second task, KL was required to construct grammatically correct sentences from printed word cards presented in random order (Carramazza & Hillis, 1989). KL performed well on the grammatical judgment task (16/18 pairs correct); however, he was unable to construct eight of the nine sentence anagrams. These findings are consistent with the above results and suggest that KL's grammatical representation was largely intact; however, he was unable to produce his own grammatical structures.

KL did not show the classic dissociation between nonpropositional and propositional speech production (see Table 2). He performed poorly on the automatic speech tasks of the BDAE, as well as a phrase completion task specifically designed by Lum and Ellis (1994) to assess this dissociation. This latter task comprises 30 phrases requiring completion, which are either familiar (nonpropositional) or unfamiliar (propositional)—for example, (1) familiar phrase, "As green as *grass*," and (2) unfamiliar phrase, "As long as *grass*." KL's performance ranked in the bottom quartile for both types of phrases indicating impairments in both automatic and propositional speech (see Table 2).

KL's performance of the Supplementary ideomotor tasks of the BDAE was symptomatic of ideomotor apraxia. He tended to verbalize actions and used a finger as the mimed object in the transitive limb tasks.

MUSICAL SKILLS

KL indicated that he listened to music daily and occasionally played the keyboard and harmonica left-handed; however, he no longer composed music. His musical abilities were assessed using the Measures of Musical Abilities (Bentley, 1985) and standard items from the aural assessment tasks of the Australian Music Examinations Board (Nickson & Black, 1962). The results showed that KL's musical abilities generally fell within the normal range; however, his discrimination and production of more complex, novel rhythmic patterns were severely impaired (see Table 3). In contrast, he could still maintain a steady beat to music and when unaccompanied.

KL also showed impaired pitch working memory as tested using Zatorre's pitch short-term memory task (Zatorre & Samson, 1991). This most likely hampered his ability to accurately reproduce more complex novel melodies, despite singing tuneful harmonies and

TABLE 2. Assessment of the nonpropositional/propositional speech dichotomy

Task	Performance	Comments
Boston Diagnostic Aphasic Examination		
Counting to 10	Unimpaired	
Expletives	Unimpaired	
Reciting days of the week	Impaired	Difficulty with "Wednesday"; "Saturday" was said as "Sunday"
Reciting the alphabet	Impaired	Only able to say "a, b, c, d"; singing the alphabet song did not facilitate performance
Reciting months of year	Impaired	
Reciting nursery rhymes[a]	Impaired	Able to say some words with continuous prompting
Phrase completion task		
Familiar phrases	12/30	25th percentile[b]
Unfamiliar phrases	6/30	25th percentile[b]

[a] Nursery rhymes included "Jack and Jill" and "Baa Baa Black Sheep."
[b] From Lum & Ellis (1994).

familiar melodies throughout the duration of the study. He demonstrated the ability to learn new songs as well as "relearn" songs he had performed pre-stroke. Interestingly, he was unable to sing the unrehearsed name in "Happy Birthday to You" and would leave a space for the listener to fill. He was not able to hold a conversation through song and often needed prompting to begin singing.

TABLE 3. Assessment of KL's musical skills

Task	Score	Norms
Measures of musical abilities		
Pitch discrimination	17/20	18.0 (± 1.9)[a]
Melody discrimination	8/10	8.8 (± 1.0)[a]
Chord discrimination	13/20	13.8 (± 2.9)[a]
Rhythm discrimination	0/10	8.3 (± 1.0)[a]
AMEB aural tasks[b]		
Novel pitch production		
Short melodies	3/5	
Long melodies	0/5	
Novel rhythm production	1/4	KL was able to tap the rhythms of highly familiar songs
Zatorre's pitch short-term memory task		
Control condition	24/24	
Interference condition	16/24	6% (± 2.5) error for healthy adults
	(33.3% error)	12% (± 2.6) error for left temporal lobectomy patients[c]

[a] Taken from Wilson & Pressing (1999) for males aged between 62 and 70 years.
[b] Australian Music Examinations Board (Nickson & Black, 1962).
[c] Data from Zatorre & Samson (1991).

Materials

Thirty phrases of three words were generated and randomly allocated to one of three groups of ten. Each phrase was printed beneath a line drawing of the last word. Some of these drawings were taken from the Boston Naming Test (Goodglass & Kaplan, 1972) while the remainder were created in a similar style. Each group had a similar number of phrases with one-, two-, three-, or four-syllable words.

Tunes were composed for each of the Group 1 (MIT) phrases. The rhythm and pitch contour of the tunes were similar to the rhythm and prosody of conversational speech (see Figure 2), although this aspect was occasionally sacrificed to make each tune unique, as suggested by Baker (2000). Group 2 (repetition) phrases were given a slightly exaggerated rhythm that approximated the rhythms of Group 1 without violating the natural rhythm of speech (see Figure 2). Group 1 and 2 phrases were recorded onto a practice compact disc (CD) by a professional singer using personal computer recording software (Sound Forge 4.5b, Sonic Foundry). The phrases were presented on the CD in the same order as they were rehearsed during therapy sessions, beginning with the easy phrases. Group 3 phrases were unrehearsed.

Procedure

Throughout testing, KL's responses were recorded using a portable digital audiotape (DAT) recorder. Baseline performance was obtained prior to the commencement of therapy, by asking KL to say the phrases from each group in response to picture prompts, written word prompts, and spoken word prompts. Therapy was

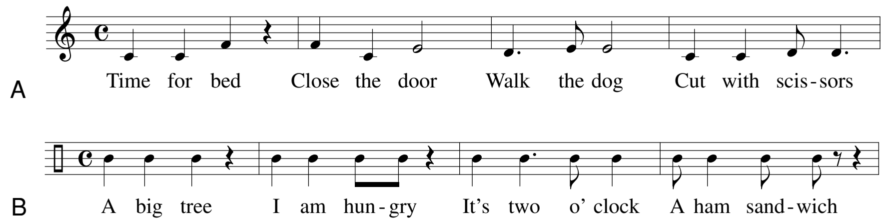


FIG. 2. Examples of the experimental stimuli. Group 1 phrases were sung while Group 2 phrases were spoken with a slightly exaggerated rhythm.

TABLE 4. Melodic Intonation Therapy procedure (Helm-Estebrooks, 1983)

Level	Group 1 phrases (MIT)	Group 2 phrases (repetition)
1	Therapist hums melody and taps rhythm	Therapist taps rhythm
2	Therapist intones phrase and taps rhythm	Therapist recites phrase and taps rhythm
3	Therapist intones phrase and taps rhythm with KL	Therapist recites phrase and taps rhythm with KL
4	Therapist intones phrase with KL, gradually fading out assistance	Therapist recites phrase with KL, gradually fading out assistance
5	KL sings phrase after therapist	KL repeats phrase after therapist
6	KL answers a question using the sung target phrase	KL answers a question using the spoken target phrase

then conducted for a period of 4 weeks by a Registered Music Therapist and the second named researcher. It entailed twice-weekly rehearsal sessions of both Group 1 and 2 phrases with the hours of training identical for each condition. Rehearsal followed the MIT method outlined by Helm-Estebrooks (1983) that includes six levels of phrase production graded according to the level of phrase difficulty and the degree of prompting provided by the therapist (see Table 4). The highest level (level 6) requires the participant to answer a question using the target phrase. KL was allowed three attempts to complete each level, and melody was not incorporated into the rehearsal of Group 2 phrases. He was encouraged to use the practice CD between sessions and his practice records showed that he did not favor one type of rehearsal over the other. Group 3 (unrehearsed) phrases were presented at baseline and at week 5, one week after the cessation of therapy (follow-up 1).

At follow-up 1, the phrases from each group were presented to KL in random order. For each phrase, he was first given a picture prompt, then a melodic prompt (for the MIT phrases only), a written word prompt, and finally a spoken or sung word prompt, until he was able to say the complete phrase. The melodic and sung word prompts used the melodies accompanying the Group 1 phrases. A similar follow-up assessment of the rehearsed phrases was conducted at week 9, five weeks after the cessation of therapy (follow-up 2). KL did not use his practice CD between follow-up 1 and 2.

A number of performance indicators were used to assess KL's phrase production. First, during therapy the highest level reached by KL for each phrase was recorded for the purpose of analyzing the number of times he reached Level 6 as a function of MIT versus repetition training. Second, his baseline and follow-up test performances were quantified using a scoring system similar to that devised by Lum and Ellis (1994). Specifically, a weighting system was used to reflect the varying difficulty of the prompts and to control for the number of syllables in each phrase (phrase length). Additionally, the production of a complete and immediately correct phrase without verbal cueing was allocated an extra two points (see Table 5). For example, an immediately correct three-syllable phrase in response to a picture prompt received a score of 8 (2 points for each syllable plus an extra 2 points for the complete phrase). The total was then divided by 6 to reflect only the use of a picture prompt to complete the phrase (see Table 5). Finally, the nature of KL's production errors was recorded to allow potential mechanisms underlying his difficulties to be examined relative to the presence or absence of training

TABLE 5. System used to score KL's performance (after Lum & Ellis, 1994)

Utterance	Score[a]	Score[b]
Correct syllable (e.g., "fish")	2	1
Approximation (e.g., "frish")	1	1/2
Complete phrase (e.g., "A cold fish")	2 + 2 + 2 + extra 2 = 8	1 + 1 + 1 = 3

Phrase length	Picture prompt	Melodic prompt	Written prompt	Spoken/sung prompt
3 syllables	score/6	score/9	score/11	score/12.5
4 syllables	score/8	score/12	score/14.67	score/16.67
5 syllables	score/10	score/15	score/18.3	score/20.8
6 syllables	score/12	score/18	score/22	score/25

[a]Without simultaneous verbal cueing.
[b]With simultaneous verbal cueing.

and the type of training received (repetition versus MIT). Of particular interest was the number of articulation errors produced by KL compared with paraphasic errors. For the purpose of coding, identifiable word approximations were classified as articulation errors. Paraphasic errors comprised word substitutions with possible phonological resemblance but no semantic resemblance to the target word. Incomplete attempts at phrases as well as the absence of an attempt were also noted.

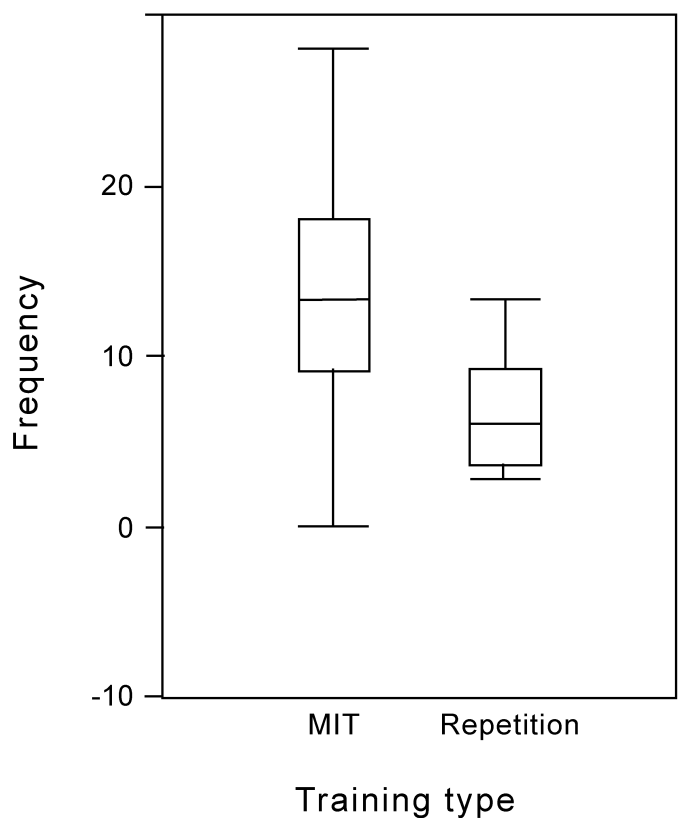


FIG. 3. Box plot showing the mean number of times KL reached Level 6 for Group 1 and 2 phrases during training. The box plot visually displays the distribution of the data, with the box representing the 25th to 75th percentile scores, and the median value indicated by a line across the box.

Results

Performance During Therapy: MIT versus Repetition Training

Exploratory data analysis revealed that the data were suitable for parametric analyses that were performed using SPSS statistical software. To examine the effects of MIT during training, a *t* test was used to compare the mean number of times KL reached Level 6 across the eight sessions of therapy as a function of the type of training. As shown in Figure 3, a significant performance advantage was evident for phrases rehearsed using MIT versus those rehearsed using repetition, $t(14) = 2.27$, $p < .05$.

Immediate Efficacy of MIT: Baseline versus Follow-up 1

A repeated measures analysis of covariance (ANCOVA) was used to examine the proportion of words correctly produced by KL as a function of time of assessment (baseline versus follow-up 1) and phrase group (1–3). Phrase length was used as the covariate in the analysis. The results showed a significant main effect of time, $F(1, 56) = 6.47$, $p < .05$, and phrase group, $F(2, 56) = 13.9$, $p < .001$, and a significant interaction between time and group, $F(2, 56) = 9.95$, $p < .001$. Phrase length had no significant effect, $F(1, 56) = 0.35$, *n.s.* As shown in Figure 4, pairwise comparisons of the estimated marginal means revealed that KL's performance of the MIT and repetition phrases was significantly better than his performance of the unrehearsed phrases across time (baseline and follow-up 1). In contrast, the difference between his overall performance of the MIT and repetition phrases was not significant.

Longer-term Efficacy of MIT: Baseline versus Follow-up 1 & 2

A repeated measures ANCOVA was used to assess the longer-term efficacy of MIT by examining the proportion of words correctly produced by KL for the

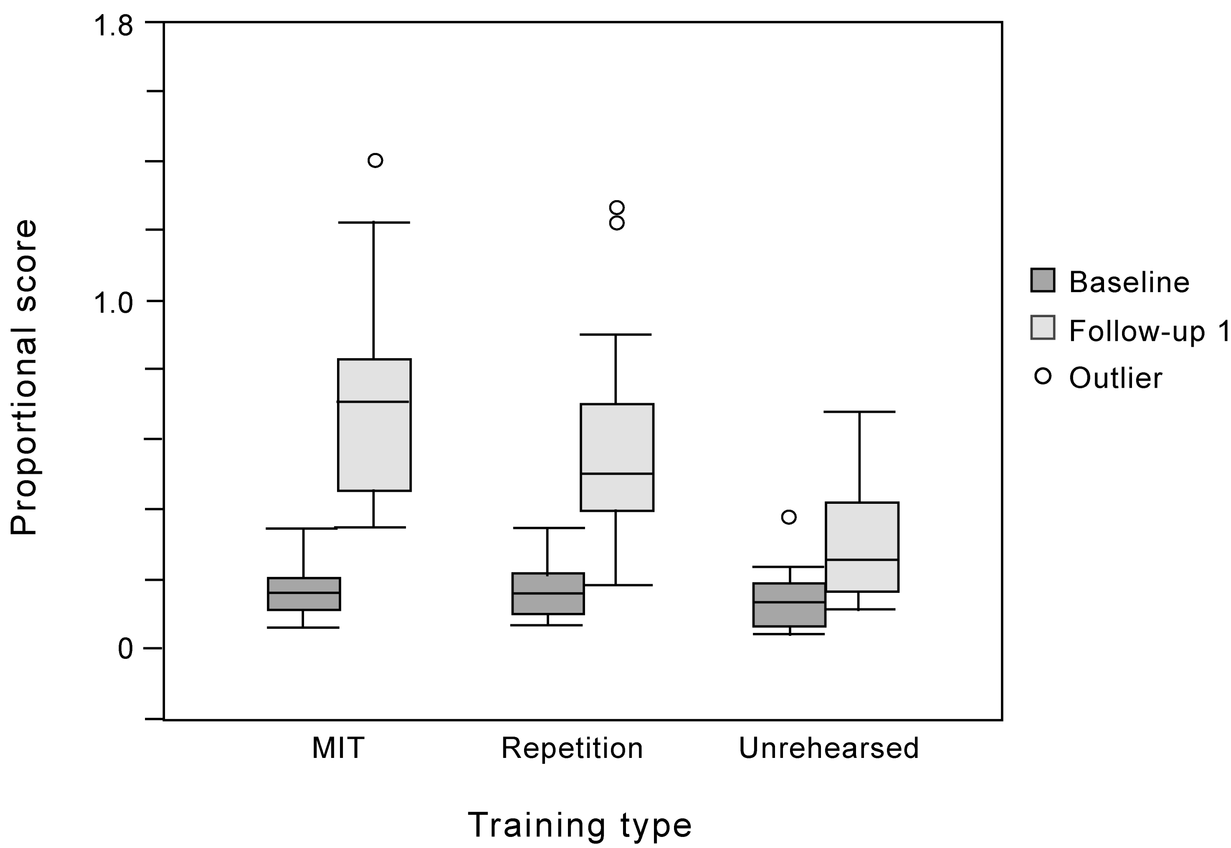


FIG. 4. Box plot showing the proportion of words correct at baseline and follow-up 1 as a function of training. Differences between the estimated marginal means collapsed across time were as follows: (1) MIT − Repetition = 0.07, *n.s.*, (2) MIT − Unrehearsed = 0.21, $p < .001$, and (3) Repetition − Unrehearsed = 0.15, $p < .01$.

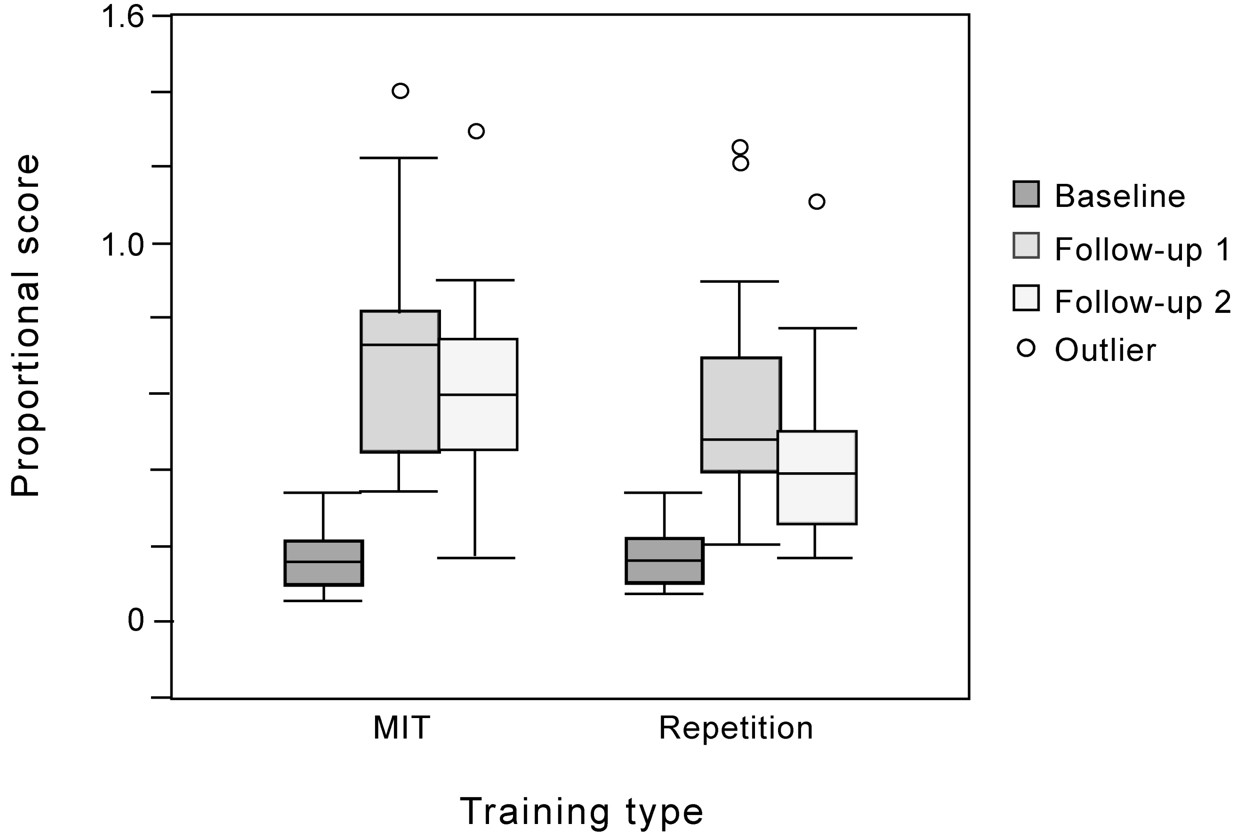


FIG. 5. Box plot showing proportion of words correct at baseline, follow-up 1 and follow-up 2 for rehearsed phrases. Parameter estimates for the MIT versus repetition phrases were as follows: (1) Baseline; B = −.002, t = −0.1, *n.s.*, (2) Follow-up 1; B = .13, t = 1.44, *n.s.*, and (3) Follow-up 2; B = .18, t = 2.3, $p < 0.05$.

rehearsed phrases (MIT versus repetition) as a function of time (baseline versus follow-up 1, 2). Phrase length was again used as the covariate in the analysis. The results showed a significant main effect of phrase group, $F(1, 37) = 5.08$, $p < .05$, and a significant interaction effect between time and group, $F(1, 37) = 5.4$, $p < .05$. In line with the previous analysis, parameter estimates indicated that there was no difference between KL's performance of the MIT and repetition phrases at baseline and follow-up 1. His performance of the repetition phrases, however, deteriorated at a faster rate than the MIT phrases, creating a significant performance advantage for the MIT phrases at follow-up 2 (see Figure 5).

TABLE 6. Breakdown of KL's production errors (%) as a function of phrase group

Performance	Group 1 (MIT)	Group 2 (repetition)	Group 3 (unrehearsed)
Correct response	8	6	–
Correct with cue	9	15	7
Incomplete phrase	19	32	26
Other MIT phrase	12	3	13
Other repetition phrase	–	6	2
Other unrehearsed phrase	–	3	5
Articulation error	34	27	34
No attempt	18	8	13

Error Analysis

Table 6 displays the types of production errors made by KL relative to the number of correct responses (with or without a prompt), shown as a function of phrase group. In all three groups, production errors included the absence of a phrase attempt, an incomplete phrase attempt, articulation errors, and paraphasic errors. Examination of the paraphasic errors revealed that they naturally subdivided according to word substitutions from other MIT phrases, other repetition phrases or other unrehearsed phrases (see Table 6). Across the three groups, the majority of errors related to articulation problems ($M = 31.67\%$), followed by incomplete phrase attempts ($M = 25.67\%$). The frequency of correct responses with or without a cue ($M = 15\%$) was similar to the number of paraphasic errors collapsed across subdivisions ($M = 14.67\%$) and the absence of phrase attempts ($M = 13\%$).

For MIT phrases, KL's paraphasic errors only consisted of other words or phrases that had been rehearsed with melody. This suggests that the melodic prompts facilitated access to words that had been rehearsed with melody over those that had not. Similarly, for repetition phrases, the majority of paraphasic errors comprised word substitutions from other repetition phrases, compared with other MIT or unrehearsed phrases. The majority of paraphasic errors in the unrehearsed group comprised other MIT phrases. This suggests that in the absence of any rehearsed articulation pattern, the MIT phrases dominated over the nonmelodic phrases. Overall, the MIT phrases showed the highest correct response rate without a prompt, whereas repetition phrases were more likely to be correct with a verbal cue. MIT was also associated with the lowest number of incomplete phrases, in contrast to the highest number for repetition phrases. Finally, unrehearsed phrases showed the lowest number of correct responses (with or without a prompt) and the highest number of paraphasic errors, supporting the general benefits of rehearsal.

Discussion

This study provides an important step in demonstrating the efficacy of MIT in a neurologically stable amateur male musician with severe expressive aphasia using a controlled pre- versus post-treatment design. In particular, during therapy KL was significantly more likely to reach the stage where he could answer a question with a sung target phrase than a spoken phrase. Although sung or spoken rehearsal had a short-term beneficial effect on his word production compared with no training, the effects of MIT were more durable, facilitating superior phrase production 5 weeks after therapy. MIT phrases were also more commonly produced without a prompt and were more likely to be complete utterances.

The lack of an effect of melody on speech production in previous research using control tasks with aphasic patients may be explained by the absence of training in these studies (cf. Cohen & Ford, 1995; Hébert et al., 2003; Peretz et al., 2004). Rehearsal was of paramount importance for KL's speech production, facilitating both melodic and nonmelodic phrases. The effect of rehearsal provides an important clue about the possible mechanisms underlying the efficacy of MIT in this case. Specifically, MIT appeared to facilitate KL's most obvious difficulty with correctly articulating speech sounds. In fact, rehearsal generally aided his articulation and sequencing of speech sounds, as evident from the immediate and beneficial effects of both MIT and repetition. Both of these strategies employed a more pronounced rhythm, which had the effect of slowing the tempo and providing greater accentuation of the speech sounds, likely reducing the complexity of the speech production process (Sparks & Holland, 1976). Interestingly, the additional use of pitch in the MIT phrases had a longer-term effect of facilitating phrase production, highlighting the importance of the melodic component in MIT.

The relative contribution of pitch and rhythm in the efficacy of MIT is an unresolved issue. Boucher and colleagues (2001) previously examined the effects of pitch versus rhythm on speech production in two aphasic patients, reporting that rhythm facilitated speech articulation to a greater degree than the tonal attributes of speech. Their paradigm, however, did not utilize MIT preventing direct comparison of their findings. KL's overall pattern of performance suggests that the tonal attributes of MIT had less impact on speech articulation and more impact on phrase encoding or access at a cognitive level. Specifically, a beneficial, articulatory effect of melody above that provided by rhythm should have been evident as superior MIT phrase performance immediately after therapy (i.e., MIT > repetition across baseline and follow-up 1). The advantage of melodic over nonmelodic rehearsal, however, appeared as a delayed recall effect (follow-up 2). This may indicate that the act of combining melody and speech promoted different storage or access to the sung phrases that became apparent once the initial benefits of rehearsal had subsided, resulting in more effective, longer-term production in the presence of expressive aphasia.

This interpretation is consistent with observations of isolable processing systems for speech and music (Hébert et al., 2003; Marin, 1982; Peretz & Coltheart, 2003; Peretz & Morais, 1993). KL's production errors are also consistent with this interpretation. Specifically, his paraphasic errors of MIT phrases only consisted of other MIT phrases, supporting separate representations or access to rehearsed sung phrases and verbally repeated phrases. Use of the melodic representation appeared to dominate in the presence of expressive aphasia, giving rise to the higher incidence of MIT paraphasias in the unrehearsed condition. In contrast, KL's representation of verbally repeated phrases appeared more difficult to access, requiring a greater degree of verbal cueing and resulting in more frequent incomplete utterances. Overall, our findings may account for the previously reported paradoxical effect in the literature of preserved singing of familiar songs in patients with expressive aphasia, but the lack of an effect for *unrehearsed* sung over spoken novel stimuli in such patients (Cohen & Ford, 1995; Hébert et al., 2003; Peretz et al, 2004; Smith, 1966; Yamadori et al., 1977). In other words, *melodic rehearsal* appeared to be the key component for effecting longer-term improvements in speech production.

The classic dissociation between nonpropositional and propositional speech production was not supported in this case, rather KL showed impairments in both domains. Four weeks of intensive MIT did not appear to facilitate KL's spontaneous propositional speech, contrary to previous reports in the literature (Albert & Sparks, 1973; Naeser & Helm-Estebrooks, 1985; Sparks et al., 1974). It is possible that 4 weeks of therapy did not allow sufficient time to observe any effects on generative speech (Baker, 2000). On the other hand, patients in previous studies have begun MIT soon after the onset of aphasia, and may have recovered generative speech regardless of therapy. Given the findings of this study, it is likely that a longer course of therapy would have allowed the formation of close associations between a larger range of melodies and phrases. This, in turn, could have augmented the number of well-rehearsed, sung phrases incorporated in daily conversations. It seems unlikely, however, that a longer course of therapy would have facilitated KL's generative speech production per se. Rather, his generative impairment appeared pervasive, affecting not only the production of grammatically correct linguistic structures but also, possibly, the ability to compose new music. The ability of MIT to affect a transfer of learning to everyday speech is an important issue that warrants further investigation through the use of an experimental paradigm of sufficient length directly targeting this issue.

Finally, the dissociation in KL's rhythmic skills was consistent with previous amusic cases, and provides support for isolable processing subsystems within the music domain (Peretz, 1993; Peretz et al., 1994; Wilson & Pressing, 1999). In particular, KL was able to maintain a steady beat when unaccompanied or in response to music, despite an impairment of more complex rhythmic pattern discrimination and reproduction. Peretz (1990; Peretz & Morais, 1993) has previously ascribed this dissociation to the role of the left hemisphere in local rhythmic functions such as grouping, while the right hemisphere mediates global features such as meter. We have recently reported the case of a musician (HJ) with the reverse pattern of rhythmic deficits following a right temporoparietal infarct (Wilson, Pressing, & Wales, 2002). Specifically, HJ showed impaired ability to maintain a steady beat despite accurate discrimination of nonmetrical rhythmic patterns. Importantly, KL's disturbance of more complex rhythmic processing did not appear to prevent adequate performance of the MIT and repetition phrases used in the current study.

While it is acknowledged that the results of this study are necessarily limited to a single case, we believe they highlight a number of generic issues in this area of research. Most notably, we have demonstrated the importance of using a pre- versus post-treatment design that accounts for confounding factors such as spontaneous recovery, to assess the efficacy of MIT. In the clinical setting, the need to use MIT is relatively infrequent, pertaining to a certain type of neurological insult with associated cognitive effects. In this regard, KL presented as an ideal candidate in which to assess the efficacy of MIT. His background as an amateur singer provided a high level of motivation for him to engage in the study. Arousal in response to music has been well documented in both musicians and nonmusicians (Crncec, Wilson, & Prior, 2006), highlighting more generally the potential benefits of music.

Author Note

The authors would like to express their gratitude to KL, whose enthusiasm and perseverance made this research possible. We would also like to recognize the important contribution of Cate Ahrendt (Registered Music Therapist) for her invaluable expertise with Melodic Intonation Therapy, and A/Prof Michael Saling for his expert advice on aphasia. This research was partly supported by funding from the National Health & Medical Research Council of Australia.

Address correspondence to: Dr. Sarah J. Wilson, School of Behavioural Science, The University of Melbourne, Victoria 3010, Australia. E-MAIL *sarahw@unimelb.edu.au*

References

Albert, M. L., Sparks, R. W., & Helm, N. A. (1973). Melodic intonation therapy for aphasia. *Archives of Neurology, 29,* 130–131.

Baker, F. (2000). Modifying the melodic intonation therapy program for adults with severe non-fluent aphasia. *Music Therapy, 18,* 110–114.

Belin, P., Eeckhout, P. V., Zilbovicius, M., Remy, P., François, C., Guillaume, S., et al. (1996). Recovery from nonfluent aphasia after melodic intonation therapy: A PET study. *Neurology, 47,* 1504–1511.

Bentley, A. (1985). *Measures of musical abilities.* London: Nfer-Nelson.

Botez, M., & Wertheim, N. (1959). Expressive aphasia and amusia following right frontal lesion in a right-handed man. *Brain, 82,* 186–202.

Boucher, V., Garcia, L. J., Fleurant, J., & Paradis, J. (2001). Variable efficacy of rhythm and tone in melody-based interventions: Implications for the assumption of a right-hemisphere facilitation in non-fluent aphasia. *Aphasiology, 15,* 131–149.

Brust, J. C. M. (1980). Music and language: Musical alexia and agraphia. *Brain, 103,* 367–392.

Brust, J. C. M. (2001). Music and the neurologist: A historical perspective. *Annals of the New York Academy of Sciences, 930,* 143–152.

Carramazza, A., & Hillis, A. E. (1989). The disruption of sentence production: Some dissociations. *Brain and Language, 36,* 625–650.

Code, C. (1997). Can the right hemisphere speak? *Brain and Language, 57,* 38–59.

Cohen, N. S., & Ford, J. (1995). The effect of musical cues on the nonpurposive speech of persons with aphasia. *Journal of Music Therapy, 32,* 46–57.

Crncec, R., Wilson, S. J., & Prior, M. (2006). The cognitive and academic benefits of music to children: Facts and fiction. *Educational Psychology*, 26, 579–594.

Donnan, G. A., Carey, L. M., & Saling, M. M. (1999). More (or less) on Broca. *The Lancet, 353,* 1031–1032.

Ellis, A. W., & Young, A. W. (1996). *Human cognitive neuropsychology: A text book with readings.* East Sussex: Psychology Press.

Epstein, C. M., Meador, K. J., Loring, D. W., Wright, R. J., Weissman, J. D., Sheppard, S., et al. (1999). Localisation and characterization of speech arrest during transcranial magnetic stimulation. *Clinical Neuropsychology, 110,* 1073–1079.

Geschwind, N., Quadfasel, F. A., & Segarra, J. M. (1968). Isolation of the speech area. *Neuropsychologia, 6,* 327–340.

Gleason, J. B., & Goodglass, H. (1984). Some neurological and linguistic accompaniments of the fluent and nonfluent aphasics. *Topics in Language Disorders, 4,* 71–81.

Goodglass, H., & Kaplan, E. (1972). *The assessment of aphasia and related disorders.* London: Henry Kimpton.

Gordon, H. W., & Bogen, J. E. (1974). Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *Journal of Neurology, Neurosurgery, and Psychiatry, 37,* 727–738.

Ham, R. E. (1990). *Therapy of stuttering.* Englewood Cliffs, NJ: Prentice Hall.

Hébert, S., Racette, A., Gagnon, L., & Peretz, I. (2003). Revisiting the dissociation between singing and speaking in expressive aphasia. *Brain, 126,* 1838–1850.

Helm-Estebrooks, N. (1983). Exploiting the right hemisphere for language rehabilitation: Melodic intonation therapy. In E. Perecman (Ed.), *Cognitive processing in the right hemisphere.* New York: Academic Press.

Jackson, J. H. (1866). Notes on the physiology and pathology of language. *Medical Times and Gazette, 1,* 659.

Jackson, J. H. (1871). Singing by speechless (aphasic) children. *The Lancet, 2,* 430–431.

Jackson, J. H. (1874). On the nature of the duality of the brain. In J. Taylor (Ed.), *Selected writings of John Hughlings Jackson* (pp. 129–145). New York: Basic Books.

Jackson, J. H. (1915). On affections of speech from disease of the brain. *Brain, 38,* 1–27.

Laughlin, S. A., Naeser, M. A., & Gordon, W. P. (1979). Effects of three syllable durations using the melodic intonation therapy technique. *Journal of Speech, Language and Hearing Research, 22,* 311–320.

Linebarger, M. C., Schwartz, M. F., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition, 13,* 361–392.

Lum, C. C., & Ellis, A. W. (1994). Is "nonpropositional" speech preserved in aphasia? *Brain and Language, 46,* 368–391.

Marin, O. S. M. (1982). Neurological aspects of music perception and performance. In D. Deutsch (Ed.), *The psychology of music* (pp. 453–477). New York: Academic Press.

Naeser, M. A., & Helm-Estebrooks, N. (1985). CT scan localisation and response to melodic intonation therapy. *Cortex, 21,* 203–223.

Nenonen, M., Niemi, J., & Laine, M. (2002). Representation and processing of idioms: Evidence from aphasia. *Journal of Neurolinguistics, 15,* 43–58.

Nickson, N., & Black, E. (1962). *Specimen aural tests for the examinations of the Australian Music Examinations Board.* Melbourne: Allan & Co.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain damaged patients. *Brain, 113,* 1185–1205.

Peretz, I. (1993). Auditory agnosia: A functional analysis. In S. McAdams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 199–230). Oxford: Clarendon Press.

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience, 6,* 1–4.

Peretz, I., Gagnon, L., Hébert, S., & Macoir, J. (2004). Singing in the brain: Insights from cognitive neuropsychology. *Music Perception, 21,* 373–390.

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., et al. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain, 117,* 1283–1301.

Peretz, I., & Morais, J. (1993). Specificity for music. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 373–390). Amsterdam: Elsevier.

Perry, D. W., Zatorre, R. J., Petrides, M., Alivisatos, B., Meyer, E., & Evans, A. C. (1999). Localization of cerebral activity during simple singing. *Neuroreport, 10,* 3979–3984.

Riecker, A., Ackermann, H., Wildgrubber, D., Dogil, G., & Grodd, W. (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport, 11,* 1997–2000.

Smith, A. (1966). Speech and other functions after left (dominant) hemispherectomy. *Journal of Neurology, Neurosurgery, and Psychiatry, 29,* 467–471.

Sparks, R., Helm, N., & Albert, M. (1974). Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex, 10,* 303–316.

Sparks, R. W., & Holland, A. L. (1976). Method: Melodic intonation therapy for aphasia. *Journal of Speech and Hearing Disorders, 41*, 287–297.

Speedie, L. J., Wertman, E., Ta'ir, J., & Heilman, K. (1993). Disruption of automatic speech following a right basal ganglia lesion. *Neurology, 43,* 1768–1774.

Stewart, L., Walsh, V., Frith, U., & Rothwell, J. (2001). Transcranial magnetic stimulation produces speech arrest but not song arrest. *Annals of the New York Academy of Sciences, 930,* 433–435.

Wechsler, D. (1987). *Wechsler Memory Scale—Revised.* San Antonio: Harcourt Brace Jovanovich.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale* (3rd ed.). San Antonio: The Psychological Corporation.

Wildgruber, D., Ackermann, H., Klose, U., Kardatzki, B., & Grodd, W. (1996). Functional lateralization of speech at primary motor cortex: A fMRI study. *Neuroreport, 7,* 2791–2795.

Wilson, S. J., & Pressing, J. (1999). Neuropsychological assessment and modelling of musical deficits. In R. R. Pratt & D. Erdonmez Grocke (Eds.), *Music medicine 3. Music medicine and music therapy: Expanding horizons* (pp. 47–74). Melbourne: University of Melbourne.

Wilson, S. J., Pressing, J. L., & Wales, R. J. (2002). Modelling rhythmic function in a musician post-stroke. *Neuropsychologia, 40,* 1494–1505.

Yamadori, A., Osumi, Y., Masuhar, S., & Okubo, M. (1977). Preservation of singing in Broca's aphasia. *Journal of Neurology, Neurosurgery, and Psychiatry, 40,* 221–224.

Zatorre, R. J., & Samson, S. (1991). Role of the right temporal neocortex in retention of pitch in auditory short-term memory. *Brain, 114,* 2403–2417.

Physiotherapy Theory and Practice, 21(4):257–269, 2005

ISSN: 0959-3985 print/1521-0510 online
DOI: 10.1080/09593980500321085

Taylor & Francis
Taylor & Francis Group

Interactive Metronome® training for a 9-year-old boy with attention and motor coordination difficulties

Melinda L. Bartscherer, PT, MS*,[1] and Robin L. Dole, PT, EdD, PCS[2]
[1]*President, Dynamic Therapy Associates, 1219 Goodman Drive, Fort Washington, Pennsylvania, 19034*
[2]*Associate Dean and Program Director at the Institute for Physical Therapy Education, Widener University, One University Place, Chester, Pennsylvania, 19013*

The purpose of this case report is to describe a new intervention, the Interactive Metronome,®[1] for improving timing and coordination. A nine-year-old boy, with difficulties in attention and developmental delay of unspecified origin underwent a seven-week training program with the Interactive Metronome.® Before, during, and after training timing, accuracy was assessed with testing procedures consistent with the Interactive Metronome® training protocol. Before and after training, his gross and fine motor skills were examined with the Bruininiks-Oseretsky Test of Motor Proficiency (BOTMP). The child exhibited marked change in scores on both timing accuracy and several BOTMP subtests. Additionally his mother relayed anecdotal reports of changes in behavior at home. This child's participation in a new intervention for improving timing and coordination was associated with changes in timing accuracy, gross and fine motor abilities, and parent reported behaviors. These findings warrant further study.

Introduction

This case report discusses the use of a new computerized intervention that is aimed at improving attention, timing, sequencing, and coordination. Initial reports indicate that this technology, the Interactive Metronome® (IM)[1], may be a useful tool in increasing attention, promoting academic skills, and decreasing aggression in young boys with attention deficit hyperactivity disorder (Shaffer et al, 2001). The intervention is gaining popularity in the media and parents of children with attention and motor coordination difficulties are seeking out individuals who are trained in the approach. Physical therapists need to be aware of the intervention and the relevant science to support it, in order to provide effective consultation and recommendations to their patients and families.

Improving timing, sequencing, and coordination is often a goal of physical therapy but can be quite difficult to accomplish. Overall, therapists have little at their disposal to draw on for guidance other than "common sense" interventions, which appear to have little more than face validity. The purpose of this case report was to describe the application of the IM Intervention on timing for one child and

Accepted for publication 14 October 2004.
Information and data in this case report, in part, was presented at the APTA Combined Sections Meeting in Boston, Massachusetts (2002).
This case report was completed in partial fulfillment of the requirements of Ms. Bartscherer's Doctor of Science degree, Rocky Mountain University of Health Professions, Provo, Utah.
*Clinical Assistant Professor, Institute for Physical Therapy Education, Widener University when this case report was written.
[1]Interactive Metronome® – 2500 Weston Road, Weston, Florida 33331.
Address correspondence to Dr. Robin L. Dole, Institute for Physical Therapy Education, Widener University, One University Place, Chester, PA 19013. E-mail: Robin.L.Dole@Widener.Edu

present the ways in which this child changed over the course of seven weeks.

Evaluating this intervention for scientific merit is important for evidence-based physical therapy practice. This report will also discuss what is currently known regarding the scientific validity of this new intervention and use this information to provide possible explanations for the outcomes reported in this case.

Case description

"John" was a nine-year-old, right-handed, African-American male, whose mother was interested in reports in the news media about a training program that claimed to help improve attention and organization. She described her son as having "difficulty concentrating" and not being "very coordinated." She related that it took him "a long time to get anything done," and that his teachers were concerned over his lack of ability to complete tests within the required timeframe. Additionally he "never liked to color, draw, play with playdough, or do cutting," but has "always been good at figuring things out that interested him." She also commented that he made loose knots in his shoes and it was difficult for him to button his pants.

During the year prior, he attended a private school that focused on experiential learning and therefore did not receive any special education or therapeutic services. At his parents' request, he had not been officially evaluated for learning difficulties, attention problems, or other developmental delays to avoid having him "labeled." He also was not taking medications during the time of this intervention. His mother reported a normal birth and a medical history negative for trauma or illnesses.

Examination

The child was a quiet, thin boy who appeared slightly small for his age. He was polite and followed directions well. During the initial session, it was apparent that he had difficulties with speech articulation that made it difficult to understand him the majority of the time. Additionally, he had difficulty coordinating the movements that comprise the initial metronome testing sequence and got easily tangled in the wires for the headphones and hand trigger that are used in both testing and training sessions. During the initial session, the Interactive Metronome® long-form test (IM LFT) and the Bruininks-Oseretsky Test of Motor Proficiency (BOTMP) (Bruininks, 1978) were administered.

Interactive metronome® program

The Interactive Metronome® is a PC-based version of a traditional music metronome and was originally developed for improving timing accuracy in musicians. The equipment utilized for this program included a laptop computer with 200 MHz Pentium Processor, Windows 98 and IMPro 5.0 software. Additional hardware included a hand trigger, foot trigger, and two stereo headphones. The hand trigger is a small circular pressure sensitive trigger approximately 1″ in diameter. It connects by Velcro to a cuff strap that attaches around the individual's hand. The foot trigger consists of two $1'' \times 10''$ trigger strips aligned in parallel and placed inside a vinyl pad. Both hand and foot triggers are connected to the computer by cables. The stereo headphones are connected to a splitter and then to the computer, thus allowing both the trainer and participant to hear the metronome beat and feedback sounds.

The IMPro 5.0 software generates a computer-based metronome at a set frequency of 54 beats per minute and, through guide sounds, provides bandwidth feedback to the learner following each practice trial (beat of the metronome). The learner performs 1 of 13 tasks and tries to activate the trigger in time with the reference beat of the metronome. The program calculates timing error (absolute error in milliseconds) for each task performed in the session. Additionally recorded for each task are: number of trials considered very early, early, late, and very late, number of trials completed, percentage of trials in which the timing error is 15 milliseconds (ms) or less, and the greatest number of trials in a row in which the timing error is 15 milliseconds or less.

Feedback on timing error, via "guide sounds," is given during training at a relative frequency of 100%. Feedback is auditory only and heard by the learner as high and low tones through stereo headphones. If the learner activates the trigger prior to the metronome beat, the guide sounds are heard in the left ear.

If trigger activation occurs following the metronome beat the guide sounds are heard in the right ear. If trigger activation is within 15 milliseconds of the metronome beat the guide sounds are heard in both ears.

Interactive metronome® testing

The IM LFT is comprised of 14 tasks. The first 13 tasks (see Table 1) are performed with the metronome reference beat only, and the last task (task number 14) is a repeat of the first task and is performed with the metronome beat and guide sounds. When the guide sounds are turned on, the child will hear a second sound that is separate from the reference beat and occurs at the time the trigger is hit. This timing of this sound indicates how close the child's trigger activation is to the metronome reference beat. The guide sounds appear in the left or right ear (indicating either being before or after the metronome beat) and with lower and higher tones (indicating how far off the beat). As the child tries to hit the trigger in time with the metronome beat, timing is measured as the number of milliseconds before or after the beat. Average milliseconds are

Table 1. Descriptions of the 13 IM tasks.

IM task	Description
Both hands	Both hands are moved in circles such that the trigger is hit in midline followed by the hands moving up, out, and around. Circles are approximately 10″ in diameter.
Right hand	The right hand makes a horizontal circle making contact with the thigh then moving forward, out, and back.
Left hand	The left hand makes a horizontal circle making contact with the thigh then moving forward, out, and back.
Both toes	One foot moves forward to tap the toes on the trigger pad while weight is supported on the other foot. Feet are alternated.
Right toe	Weight is supported on the left foot while the right foot toe taps on the trigger pad. The heel of the tapping foot remains on the floor.
Left toe	Weight is supported on the right foot while the left foot toe taps on the trigger pad. The heel of the tapping foot remains on the floor.
Both heels	The trigger pad is placed behind the individual. The individual alternately taps heels on the trigger pad.
Right heel	Weight is supported on the left foot while the right heel taps on the trigger pad. The toe of the tapping foot remains on the floor.
Left heel	Weight is supported on the right foot while the left heel taps on the trigger pad. The toe of the tapping foot remains on the floor.
Right hand/left toe	This is a combination of right hand alternated with left toe. Thus, the hand movement is slowed to occur over two beats (every other beat) as is the left toe.
Left hand/right toe	This is a combination of left hand alternated with right toe. Thus, the hand movement is slowed to occur over two beats (every other beat), as is the left toe.
Balance right foot/tap left	The individual balances on the right foot while keeping the left foot in the air and tapping the left toe on the trigger pad.
Balance left foot/tap right	The individual balances on the left foot while keeping the right foot in the air and tapping the right toe on the trigger pad.

calculated for each of the tasks. Upper limb average is calculated with those activities involving the arms and lower limb average with those activities involving the legs, including tasks that combine upper and lower limbs. Sufficient test-retest reliability ($r = .85$ to $.97$) for the IM LFT has been reported (Cassily and Jacokes, 2001).

The results of the initial IM LFT revealed that the child had significant timing and movement coordination difficulties. Table 2 describes his timing accuracy for each of the IM LFT items and the calculated upper limb, lower limb, and combined averages. The average time off of the beat at the initial LFT was 159.44 ms. According to the Interactive Metronome Indicator Chart (2003), the average pre-test performance on the long-form test for children between 9 and 10 years of age is in the range of 55–79 ms. Upon closer examination, it was noted that he was less accurate with his feet than with his hands. Additionally, those tasks requiring opposite upper and lower limb coordination were much less accurate than those relying on just one limb or alternating bilateral upper or lower limbs.

Qualitatively, the child exhibited timing deficiency patterns of the disassociative and hyperballistic types. A disassociative pattern is one in which there is no clear association between the child's response of hitting the target and the beat of the metronome. Rather, the responses appear chaotic and random without a discernable pattern. At times the responses are very early, sometimes late, sometimes the individual will respond several times in between beats, other times a beat will be totally missed. In this situation, the individual has difficulty interpreting the metronome beat in order to synchronize his or her movements to it. This difficulty led to the inability to calculate a millisecond average for the first task during his initial long form test (LFT; see Table 2). Additionally, he exhibited movements, which were hyperballistic; rather than being of uniform and smooth speed throughout the movement pattern, his movements were first slow and then very fast, forceful, and ballistic as he moved to hit the trigger. These two patterns are among the six deficiency pattern's that can be identified with the IM testing and training protocol (Burpee et al, 2001; Shaffer

Table 2. Timing accuracy (milliseconds off of the beat) on IM LFT–initial, midterm, and final sessions.

IM task	IM LFT initial (ms)	IM LFT midterm (ms)	IM LFT final (ms)
Both hands	No score obtained*	47.06	23.32
Right hand	53.37	27.38	34.50
Left hand	140.00	33.92	24.86
Both toes	76.11	75.11	26.19
Right toe	337.67	55.47	25.30
Left toe	129.93	51.04	28.07
Both heels	49.48	63.70	26.08
Right heel	69.11	85.10	61.19
Left heel	135.03	75.63	34.97
Right hand/left toe	275.00	78.37	43.88
Left hand/right toe	229.22	24.47	25.93
Balance right foot	193.79	105.35	40.90
Balance left foot	311.56	61.93	21.87
Both hands with guide sounds	110.61	39.13	24.40
Hands average	138.19	36.87	26.77
Feet average	180.69	67.62	33.44
LFT average	159.44	52.25	30.11
Normal for age	55–79		

*Child did not perform enough repetitions in the allotted time to determine a score for this task.

et al, 2001). These represent various timing difficulties that can be exhibited qualitatively and quantitatively with the program, and have been termed dissociative, contraphasic, hyper-ballistic, hyper-anticipatory, hypo-anticipatory, and auditory hypersensitivity.

The child also exhibited motor planning difficulties, which were worse with his legs than his arms. When doing the tasks of alternating toe taps, unilateral toe taps, or alternating heel taps, he had difficulty organizing the movement without hand over foot assistance. His foot would drift out of the target field or he would spontaneously shift to a different task. He was unable to continue the task pattern for more than 3 to 5 repetitions without assistance.

Motor skill testing

The child's motor skills were assessed via administration of the Bruininks-Oseretsky Test of Motor Proficiency (BOTMP). The BOTMP is a standardized, norm referenced test used to assess gross and fine motor skills in children between the ages of 4.5 and 14.5 years of age. The purported uses of the test include making decisions about educational placement, assessing gross and fine motor skills, developing and evaluating motor training programs, and assisting researchers and clinicians. Psychometric properties have been reported with established face validity, content validity, internal consistency, test-retest ($r = .88$), and inter-rater reliability ($r = .98$; Bruininks, 1978). The test is divided into gross motor, upper-limb coordination, and fine motor subtests. Gross and fine motor as well as total battery composites are calculated by converting the raw scores to standard scores based on the child's age. Percentile ranks and stanines can then be determined.

The results of the BOTMP revealed that John exhibited significant delays in both gross and fine motor skills, with the greater deficiencies noted in fine motor skills. The Battery Composite revealed that he scored below the 1st percentile rank and in the first stanine when compared to children of similar age (see Table 3).

Evaluation

According to the Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition

Table 3. Results of Bruininiks-Oseretsky test of motor proficiency - initial and final sessions.

BOTMP complete battery	Point score		Standard score		Composite score		Percentile rank		Stanine	
	Initial	Final	Initial	Final	Initial	Final	Initial	Final	Initial	Final
Gross motor subtests										
Running speed and agility	8/15	8/15	10	10						
Balance	8/32	17/32	1	3						
Bilateral coordination	10/20	11/20	13	14						
Strength	16/42	17/42	10	11						
Gross motor composite			34	38	31	34	3	6	1	2
Upper limb coordination	15/21	12/21	8	2						
Fine motor subtests										
Response speed	9/17	13/17	15	23						
Visual-motor control	12/24	13/24	2	3						
Upper limb speed and dexterity	26/72	32/72	2	7						
Fine motor composite			19	33	25	39	1	14	1	3
Battery composite			61	73	24	30	<1	2	1	1

(American Psychological Association, 1994), the child exhibited many of the behavioral characteristics consistent with two specific diagnoses; attention deficit hyperactivity disorder—inattentive type and developmental coordination disorder. Though he had not been formally diagnosed with either of these conditions, the result of motor skill testing revealed that his performance was markedly below that of his same aged peers and his parents confirmed that his inattentiveness, poor motor coordination, and impaired motor skills were negatively impacting his participation at school and home.

Diagnosis and prognosis

According to the Guide to Physical Therapist Practice, the child's difficulties are consistent with the preferred practice pattern 5B: Impaired Neuromotor Development (American Physical Therapy Association, 2001). Given that his motor skills were so far below his peers and acknowledging what is known from the literature on children with developmental coordination disorder (Barnhart, Davenport, Ebbs, and Nordquist, 2003), his difficulties are not likely to improve without intervention and can be expected to contribute to problems of low self-esteem and poor social interactions. The child would be a candidate for traditional physical therapy intervention aimed at improving his overall motor coordination and specific skills needed for active participation with his family and peers. The Guide provides an estimate of number of visits between 6 and 90 in order to achieve the expected and desired outcomes of intervention. Based on the examination findings and the information found in the literature (e.g. Pless, Carlsson, Sundelin, and Persson, 2001; Polatajko et al, 1995), it would be reasonable to expect an initial program of intervention to cover at least 10 to 20 sessions.

Intervention

The Interactive Metronome® intervention took place over a 7-week period during the summer when the child was off from school. During this time, he did not receive other therapy services nor did he participate in any school, academic or sports related activities, other than what he did at home with his family.

The Interactive Metronome® intervention consists of approximately 15 to 20 sessions of practice using a variety of upper and lower limb tasks as are found in the IM LFT. Sessions were scheduled for three times weekly, with at least one day of no training in between sessions. During the first half of the program, the primary goals were to learn the tasks and the meaning of the guide sounds that were used to provide feedback on performance. The focus of the last half of the program was on increasing accuracy and consistency. Within the program, the child optimally works up from being able to do 100 to 200 repetitions of a single task without rest to a maximum of 2,000 repetitions of a single task without rest. Repetitions are set at a frequency of 54 beats per minute; therefore a 2,000 repetition task takes approximately 40 minutes to complete.

Each session began with a retention test, the short form test (IM SFT), which consists of 2 trials of 54 repetitions (1 minute) each of the "both hands" task. For the both hands task, John moved both hands in a circular fashion in front of his body and activated the hand trigger when his hands came together. For the first trial in the IM SFT, the child heard only the metronome reference beat. In the second trial, he heard the metronome reference beat and the guide sounds. After the IM SFT, he engaged in several tasks with the total number of repetitions in a single session ranging from 1,500 to 2,500. Each session lasted approximately 1 to 1.5 hours.

In order to help decrease the dissociative timing pattern and improve his ability to perform the tasks required of the IM program, the child participated in four recommended pre-sessions prior to starting the typical 15 session IM protocol. The pre-session training was to aid in understanding the meaning of the metronome beats and guide sounds and developing basic skill at hitting the trigger in time with the metronome beat. The number of pre-sessions was not predetermined, rather pre-sessions were continued until his dissociative pattern had decreased and his understanding of the connection of his movements to hit the trigger and the metronome beat had improved. Tasks practiced during the pre-sessions included providing hand-over-hand

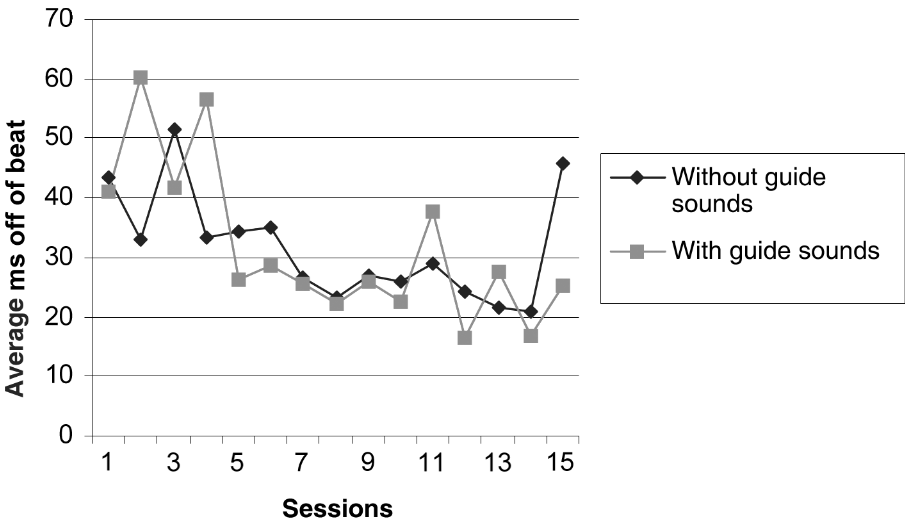


Figure 1. Short Form Test scores over time.

assistance, playing "patty cake," tapping knees then clapping hands, "high fives," and using verbalizations on the beat, in addition to the regular tasks of "both hands" and "both toes." On the fourth pre-session, the child had decreased his short form test score to 43.43 seconds without feedback sounds and 40.98 seconds with feedback sounds and it was determined that he was ready to complete the 15 session program (see Figure 1).

As the child progressed through his sessions, several items were recorded for each task practiced including the number of "right on" repetitions in a row (IAR) and the number of "bursts." Being "right on" the beat is defined as being within 15 milliseconds of the beat, with four "right ons" in a row considered a burst. These are determined by the IM software and tracked on a spreadsheet through each session. He was encouraged to do his best with the present goals relating to the number of IARs and bursts he was working toward. In general, the goals were based on what was achieved in previous sessions and his motivation. Personal best scores for bursts and IARs were recorded (see Table 4).

Example training session

All sessions were conducted in a closed and quiet room. The selection of tasks and repetitions to perform per task in a given session is prescribed in the standard IM protocol. The computer was set up on a table and both John and the therapist wore headphones. The therapist could hear the same sounds and feedback that the child heard while he practiced each task.

Table 4. Personal best records during training.

Session number	Burst record	IAR record	Task	Number of repetitions in task
2	5		Right hand	300
4	13	11	Both hands	1,000
5	19		Both hands	1,200
7	41		Both hands	1,000
9	65		Both hands	2,000
10	72	12	Both hands	1,500

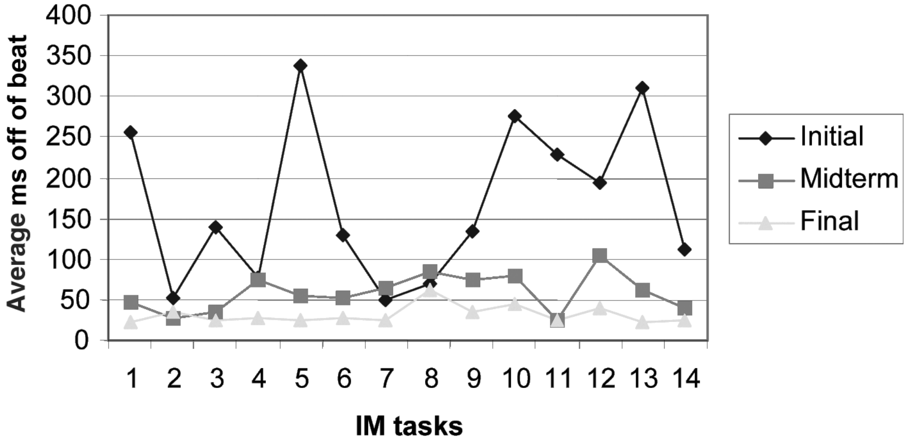


Figure 2. Performance on each task of the Long Form Test—initial, midterm, and final.

He stood for all tasks and was allowed to sit in between tasks to rest. The child had difficulty managing the cables for the hand trigger and headphones coming from the computer. He frequently became entangled such that he had difficulty doing the tasks. To improve this, he faced away from the computer so that all cables went away from his back going to the computer. Because of his difficulty with his foot movement, the floor trigger tended to move. This problem was solved by placing Velcro on the bottom of the trigger, keeping it firmly in place on the carpeted floor.

All training sessions began with the IM SFT. A verbal description and visual demonstration of the task to be performed was given prior to beginning each task. Each session followed the protocol for use with the Interactive Metronome® and contained many of the same tasks that were included in the IM LFT. Initially, physical assistance was given to his feet. This was diminished over time such that by the 6th session he no longer needed physical guidance. For a description of how the program increases the amount of time spent on a single task over the course of several sessions and the number of repetitions per session, see Table 5.

Outcomes

IM goals

Following four pre-sessions and 15 regular sessions of IM, the child had achieved many of the goals related to successfully completing the IM training protocol. He was able to perform a 2000 repetition task that required him to stay focused and moving to the metronome beat for 40 minutes. He achieved a high of 72 bursts and a high of 12 IARs ("right on" hits of the trigger in a row), each within a single 1000 or more repetition task. He was able to pay attention to the task such that he consistently corrected his pattern within one to three beats of the metronome. The child no longer required physical guidance to perform any of the movements. Qualitatively, his movements appeared much more coordinated and he had developed a new smoothness to his movements reflective of his less ballistic tendencies.

Short form tests

The child had consistent decreases in average milliseconds off the beat as seen by his short form test results (see Figure 1). This figure represents his averages during the 15-session protocol. He began at session one with SFT results in the 40-millisecond range and was able to steadily decrease these to the 20-millisecond range over the 15 sessions.

He was able to achieve many personal bests with regard to number of IARs and bursts throughout the training sessions. Table 4 represents each new record. He made quick and consistent gains in the first ten sessions. Though he continued to do well in the last 5 sessions he was

Table 5. Session descriptions to illustrate how time and repetitions on task was increased.

Session number	Number of repetitions	Description of session
Session 2	Approx. 2,400	• IM SFT (108 repetitions) • 500 repetitions of correcting faulty timing (purposefully going before or after the beat) • 300 repetitions each of correcting faulty timing in both toes, right hand, left hand, both toes, right toe, and left toe (total – 1,800)
Session 6	Approx. 2,400	• IM SFT (108 repetitions) • 500 repetitions of staying before or after the beat • 1000 repetitions of the child's choice • 100 repetitions each of balance on one foot, tap with the other (total – 200) • 300 repetitions each of right and left heel (total – 600)
Session 10	Approx. 2,500	• IM SFT (108 repetitions) • 1500 repetitions of non-leg choice • 300 repetitions each of improving the worst task from the midterm long form test, both heels, and both toes (total – 900)
Session 13	Approx. 2,500	• IM SFT (108 repetitions) • 2000 repetitions of the child's choice • 200 repetitions each of the two worst tasks from the midterm long form tests (total – 400)

not able to top the records set on session ten. As can be seen, all of his records were set with upper extremity tasks performed over relatively large numbers of repetitions (1000+ repetitions).

Long form tests

As noted in Table 2, the child made notable gains in his timing accuracy over the course of training. He began IM training with an overall timing accuracy of 159.44 ms off of the beat. By the midterm IM LFT (session 7 in the 15 session protocol) he had decreased this to 52.5 ms and by the final IM LFT had decreased this to 30.11 milliseconds. Based on the Interactive Metronome Indicator Chart (2003), for individuals of his age these scores represent a change from approximating a severe timing deficiency (160–259 ms) to being above average for his age (below 37 ms). The areas of his largest improvements included tasks with his left hand, right toe, combining hand on one side with foot on other, and balancing on one foot while tapping with the opposite toe. As seen in Figure 2, the child showed a trend toward decreased variability across the 14 tasks when comparing the initial, midterm, and final long form tests.

Motor skills testing

With regard to his performance on the BOTMP, the child again made appreciable changes over the course of seven weeks. As can be seen in Table 3, he made gains on standard scores in every subtest except for running speed and agility, and upper limb coordination. Most notably, however, are the improvements in response speed, visual motor control, and upper limb speed and dexterity. Of clinical significance is the change in percentile rank (a measure of how his skills compare to his same age peers) seen in the fine and gross motor composites as well as the battery composite. The child improved in the gross motor composite from performance in the 3rd percentile to the 6th percentile. In the fine motor composite, he improved from the 1st percentile to the 14th percentile. Lastly, he improved in the battery composite from the below the 1st percentile to the 2nd percentile. This is important clinically when considering he did not engage in any specific therapeutic interventions or recreational activities geared at improving the skills tested in the BOTMP.

Also worthy of note is the change seen in point scores on the BOTMP. In their study of the use of the BOTMP, Wilson, Polatajko, Kaplan, and Farris (1995) recommend that if this test is being used to evaluate change over time, an assessment of changes in point scores may be more reflective of change than standard and test composite scores. This provides an evaluation of the child's current performance with previous performance rather than the child's current performance relative to children from the normative sample, which would require that the child show improvements that obtained quicker than what is obtained through time and maturation. A review of the child's performance in point scores on the BOTMP (see Table 3) shows that he made the largest improvements in balance, response speed, visual-motor control, and upper limb speed and dexterity. Smaller changes were noted in bilateral coordination and strength. No change was noted in running speed and agility, and a decline in performance was noted with upper limb coordination.

Parent observations

During the second half of the training sessions, the child's parents reported changes they had noticed in his behavior. His father noted that he was more cooperative with his sister evidenced by decreased resistance to sharing TV time and in choosing sides in the car. His mother described a rule in their house that requires him to read to his younger sister everyday. Usually this turned into an argument or a fight very quickly. During the latter part of the IM training his mother noticed that the fighting had stopped and that he was now reading to his sister for 15 to 20 minutes without difficulty.

She also reported that he appeared more willing to take risks. This was evidenced to her when he asked her if he could try riding his father's bicycle and then asked for a similar "big bike" for his upcoming birthday. This was surprising to her because he had many previous, unsuccessful attempts at trying to learn to ride, and she thought he had given up. She was also surprised because his father's bike had thumb gears, which he was now able to manage.

He had been practicing math problems, as this was an area identified by his teacher as being problematic in school as he was unable to complete a series of math problems in a timely manner. His mother felt that his speed had picked up significantly during the course of IM training. She also noted that his handwriting was age-appropriate after the intervention training. She described that it had been difficult to distinguish lower from upper case letters and that he typically made all of his letters the same size. His mother reported that there was now greater distinction between the lower and uppercase letters.

Discussion

This case report discussed the outcomes related to timing and motor coordination in a 9-year-old boy after participation in a 7-week program of training with a specialized computer program using a metronome beat and guided feedback. Several notable changes occurred that might be attributable to the intervention, including the changes observed on the IM LFT, IM SFT and the BOTMP. Also of interest, however, are the anecdotal changes reported by the child's parents, some of which were related to motor function but most of which were related to affective or organizational behavior.

The literature related to the phenomenon of timing itself, specifically related to finger tapping (Ivry, 1996), and the neuroanatomical correlates related to timing (Rao et al, 1997) may be helpful in explaining the potential effect of the Interactive Metronome® training. Irvy (1996) and others (Ivry and Hazelton, 1995; Ivry and Keele, 1986) have studied the phenomenon of timing for several decades. In a frequently cited study, the results led to the hypothesis that there may be a common and central timing mechanism that governs all movements (Ivry and Hazelton, 1995). This "central clock" may be responsible for the breadth of timing issues ranging from perception of time, such as having a sense for how long a minute is or being able to distinguish between music that has a fast tempo versus a slow tempo, to being able to precisely time one's movements to an external source such as an orchestra conductor or precisely time one's agonists to antagonists in order to reach out for a glass of water. A main tenet of the theory of a central clock is that if one could find a means for training the central clock then the timing and coordination of all movements may improve. There is evidence that children considered to be "clumsy" may have altered time perception (Williams, Woollacott, and Ivry, 1992). While the child was not officially diagnosed as being clumsy or having developmental coordination disorder, he presented with many of the typical motor behaviors of children with this diagnosis. He also exhibited initial scores on the BOTMP that indicated significant delays in gross and fine motor skills.

Similar to the IM training process, experimental studies of timing of movement at the functional level using a metronome have been done to train gait in adults with central nervous system dysfunction such as stroke (Prassas, Thaut, McIntosh, and Rice, 1997) and Parkinson's Disease (McIntosh, Brown, Rice, and Thaut, 1997; McIntosh, Rice, Hurt, and Thaut, 1998). This work has shown a positive training and carryover effect with metronome training. In a training protocol that lasted three weeks, changes in temporal and kinematic components of gait were noted immediately after the training and continued through a one-month follow up (McIntosh, Rice, Hurt, and Thaut, 1998; Thaut, Kenyon, Schauer, and McIntosh, 1999).

This child was able to synchronize his movements to the beat of the metronome during the training protocol. Unlike the studies by Thaut and colleagues (Thaut, Kenyon, Schauer, and McIntosh, 1999; Thaut, McIntosh, Rice, and Prassas, 1993), however, the synchronization was not immediate. In the present case, the child required significant physical assistance and guidance to entrain his movements to the beat. One difference may lie in the presence or absence of a timing deficiency pattern and if one is present, the type of pattern. In this case report, the child exhibited a disassociative pattern that may have represented a lack of an accurate internal time precept.

The studies that used metronome training for gait were done without guide sounds or feedback and the training was task-specific. In this case, the child received guide sounds that provided feedback on accuracy. In addition the movements practiced during training were the same as those on the long and short form tests, but they were not the same as those tested on the BOTMP. It is logical to assume that practicing thousands of repetitions of a task would lead to improved performance on those tasks when a retention test without feedback is given. Evidence in the motor learning literature supports this notion (Schmidt and Lee, 1999; Shumway-Cook and Woollacott, 2001). Findings less explainable are the improvements observed in the BOTMP. The child did not practice the BOTMP test items, was not engaged in therapy interventions for motor skills, and was not involved in additional recreational activities during the weeks of IM training. If it were true that the IM training was associated with the change in BOTMP scores, this may lend support for the central clock theory. The IM training may in some way impact the function of a central time keeper such that when it is called into action for even a small set of movements and functions, overall movement and function are improved.

The child presented here exhibited some behavioral changes that where not strictly related to motor skills. His parents reported changes in his risk taking, ability to get along with his sister, ability to put his own desires aside for the sake of his sister, and his ability to stay focused to read and complete timed math problems. The connection of these behaviors to IM training appears, at the surface to be more remote. An experimental study examining the effects of IM on academic, affective, and motor performance in children with ADHD showed significant changes in the area of affective control. The children who received IM training were also significantly less aggressive (Shaffer et al, 2001) following 15 sessions of IM.

Alternative explanations

This case report was not a controlled study and as such, cannot assert that IM training contributed to the observed changes in this specific child. It is possible that other factors played a role. This child had the opportunity to work on isolated motor skills with feedback for approximately 20, one to one, 30 minute sessions. This could have accounted for some improvement in movement performance. Additionally, he received one to one instruction and attention by an adult. This may have been highly motivating to this child and contributed to the observed changes.

The child's parents were highly vested in seeing positive change in their son. Though the training was provided free of charge, they were making significant efforts to bring their son to the training and had sought out the training specifically for the types of changes noted. Their vested interest in observing change may have contributed not only to their positive interpretation of his actions but their interest may also have impacted John's behaviors.

Lastly, IM training itself is complex and multifaceted. The training program includes approximately 35,000 repetitions. Instantaneous feedback is provided on 100% of the practice trials, and there are 13 specific tasks practiced. One could hypothesize that any one of these variables may have led to the behavioral outcomes; for example, the shear number of repetitions of a task could lead to a practice effect. Further investigation is needed to understand how each of the IM components may contribute, individually and collectively, to improved timing, coordination and attention.

Conclusion

This case report provides clinical evidence that this intervention can be applied safely, and was well tolerated by the child. It also appeared to be associated with positive changes in behaviors as reported by parents and as evidenced in clinical measures. The results of this case report also raise intriguing questions that warrant future research that is well designed and controlled. Such research would help to clarify the relationship between IM training and behavioral changes including those that are motoric, affective, and organizational. If a relationship can be established, the next step may be to identify whether there are key components at play or whether the combination of IM factors is the critical aspect. In addition to forging an improved understanding of the effects of IM in children such as the child presented in this case, it will also be important to determine if a training effect occurs in other populations.

Acknowledgements

The authors wish to thank the child and his family for their participation and Widener University for the use of the Interactive Metronome® equipment and for the space to conduct the sessions.

References

American Physical Therapy Association 2001 Guide to Physical Therapist Practice, 2nd edn. Alexandria, VA

American Psychological Association 1994 Diagnostic and statistical manual of mental disorders, 4th edn. Washington, DC, Author

Barnhart RC, Davenport MJ, Ebbs SB, Nordquist VM 2003 Developmental coordination disorder. Physical Therapy 83: 722–731

Bruininks R 1978 Bruininks-Oseretsky test of motor proficiency: examiners manual. Circle Pines, MN, American Guidance Service, Inc

Burpee J, DeJean V, Frick S, Kawar M, Koomar J, Murphy Fisher D 2001 Theoretical and clinical perspectives on the Interactive Metronome (IM): a view from clinical occupational therapy practice. The American Journal of Occupational Therapy 2: 163–166

Cassily JF, Jacokes LE 2001, April 29 The Interactive Metronome®: a new computer-based technology to measure and improve timing, rhythmicity, motor planning, sequencing and cognitive capabilities. Paper presented at The Infancy and Early Childhood Training Course, Advanced Clinical Seminar, Arlington, Virginia

Interactive Metronome Indicator Chart 2003s Interactive Metronome Certified Provider Update and Resource – Part 2. Weston, FL, Interactive Metronome

Ivry RB 1996 The representation of temporal information in perception and motor control. Current Opinion in Neurobiology 6: 851–857

Ivry R, Hazelton R 1995 Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism. Journal of Experimental Psychology: Human Perception and Performance 21: 3–18

Ivry R, Keele S 1986 Timing functions of the cerebellum. Journal of Cognitive Neurosciences 6: 136–152

McIntosh G, Brown S, Rice R, Thaut M 1997 Rhythmic auditory-motor facilitation of gait patterns in patients with Parkinson's disease. Journal of Neurology, Neurosurgery, and Psychiatry 62: 22–26

McIntosh G, Rice R, Hurt C, Thaut M 1998 Long-term training effects of rhythmic auditory stimulation on gait in patients with Parkinson's disease. Movement Disorders 13: 212

Pless M, Carlsson M, Sundelin C, Persson K 2001 Pre-school children with developmental coordination disorder: self-perceived competence and group motor skill intervention. Acta Paediatric 90: 532–538

Polatajko HJ, Macnab JJ, Anstett B, Malloy-Miller T, Murphy K, Noh S 1995 A clinical trial of the process-oriented treatment approach for children with developmental coordination disorder. Developmental Medicine and Child Neurology 37: 310–319

Prassas S, Thaut M, McIntosh G, Rice R 1997 Effect of auditory rhythmic cuing on gait kinematic parameters in hemiparetic stroke patients. Gait and Posture 6: 218–223

Rao S, Harrington D, Haaland K, Bobholz J, Cox R, Binder J 1997 Distributed neural systems underlying the timing of movements. Journal of Neuroscience 17: 5528–5535

Schmidt R, Lee T 1999 Motor control and learning: a behavioral emphasis, 3rd ed. Champaign, IL, Human Kinetics

Shaffer R, Jacokes L, Cassily J, Greenspan S, Tuchman R, Stemmer P 2001 Effect of interactive metronome training on children with ADHD. American Journal of Occupational Therapy 55: 155–162

Shumway-Cook A, Woollacott M 2001 Motor control: theory and practical applications, 2nd ed. Philadelphia, Lippincott Williams & Wilkins

Thaut M, Kenyon G, Schauer M, McIntosh G 1999 The connection between rhythmicity and brain function. IEEE Engineering in Medicine and Biology March/April: 101–108

Thaut M, McIntosh G, Rice R, Prassas S 1993 Effect of auditory rhythmic cuing on temporal stride parameters and EMG patterns in hemiparetic gait of stroke patients. Journal of Neurological Rehabilitation 7: 9–16

Williams H, Woollacott M, Ivry R 1992 Timing and motor control in clumsy children. Journal of Motor Behavior 24: 165–172

Wilson BN, Polatajko HJ, Kaplan BJ, Faris P 1995 Use of the Bruninks-Oseretsky test of motor proficiency in occupational therapy. American Journal of Occupational Therapy 49: 8–17

The Examination of Upper Extremity Movement Smoothness following Stroke post Intervention with the Interactive Metronome®

Kelley Gleason, OTS

Introduction

- What is the Interactive Metronome ® IM?
- Post CVA study
- Results
- Conclusions
- Significance of study

What is the Interactive Metronome® (IM)?

- Computer-based metronome that measures and attempts to improve a person's rhythm and timing of movements
- Development of IM was based on the concept of the musical metronome
- Person attempts to perform a variety of gross motor movements in time with a reference tone
- Person's movement input is sent to the computer through hand and foot sensors

Interactive Metronome®, 2008

Hand sensor

Foot sensor

What areas does the IM impact?

- Coordination
- Attention & concentration
- Motor planning & sequencing of movements
- Word retrieval
- Language processing
- Memory
- ADL performance

Interactive Metronome®, 2006, 2008

Who benefits from IM treatment?

- Cerebrovascular Accident (CVA)
- Traumatic Brain Injury (TBI)
- Parkinson's disease (PD)
- Multiple sclerosis (MS)
- Attention deficit – hyperactivity disorder (ADHD)
- Academic performance
- Sports/athletic performance
- Cognitive deficits

Interactive Metronome®, 2008; Shaffer, 2001; Libkuman, 2002

Post CVA Study– purpose

- Experimental pre-test/post-test control group design
- Purpose – to examine the effectiveness of using the IM as a treatment modality for improving upper extremity movement smoothness following stroke
- Comparison – overall effects of IM treatment will be compared with effects of a self range of motion home exercise program

Criteria– participants

- > 3 months post stroke
- Current participants
  - Age range – 56-69
  - Average age – 61
- 3 male, 3 female
- 3 experimental, 3 control

Methods – instrumentation

- Jebsen Test of Hand Function: measures functional hand use
- 9-Hole Peg Test: measures finger dexterity and timing of fine motor movements
- Interactive Metronome® Long-form Assessment – 14 subtasks measuring rhythm and timing of movements

Hackel et al., 1992; Interactive Metronome®, 2006

Methods – procedure

- Experimental group – 8 session Interactive Metronome® treatment protocol
  - Sessions 1 & 8: Pre/ Post testing, respectively
  - Sessions 2-7: Exercise protocol

Interactive Metronome® 8 Session Treatment Protocol

Session 1	Session 2	Session 3	Session 4	Session 5	Session 6	Session 7	Session 8
Long Form Assessment	Both Hands (350 reps)	Both Hands (150 reps)	Both Hands (350 reps)	Both Hands (150 reps)	Both Hands (350 reps)	Both Hands (250 reps)	Long Form Assessment
Jebsen Test of Hand Function	Right Hand (200 reps)	Right Hand (200 reps)	Right Hand (200 reps)	Right Hand (200 reps)	Right Hand (150 reps)	Right Hand (150 reps)	Jebsen Test of Hand Function
9 – Hole Peg Test	Left Hand (200 reps)	Left Hand (200 reps)	Left Hand (200 reps)	Left Hand (200 reps)	Left Hand (150 reps)	Left Hand (150 reps)	9 – Hole Peg Test
	Right Toe (200 reps)	Right Toe (250 reps)	Right Toe (200 reps)	Right Toe (250 reps)	Right Toe (200 reps)	Right Toe (250 reps)	
	Right Heel (200 reps)	Right Heel (250 reps)	Right Heel (200 reps)	Right Heel (250 reps)	Right Heel (200 reps)	Right Heel (250 reps)	
	Left Heel (200 reps)	Left Heel (250 reps)	Left Heel (200 reps)	Left Heel (250 reps)	Left Heel (200 reps)	Left Heel (250 reps)	
	Both Hands (150 reps)	Both Hands (150 reps)	Both Hands (150 reps)	Both Hands (150 reps)	Both Hands (250 reps)	Both Hands (150 reps)	

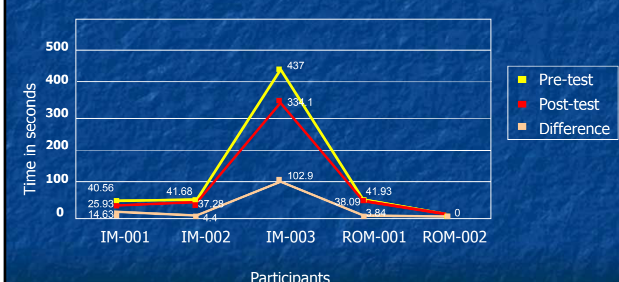
Methods - procedure

- Control group – self-range of motion home exercise program
  - Pre-test
  - 2 week ROM exercise program
  - Post-test

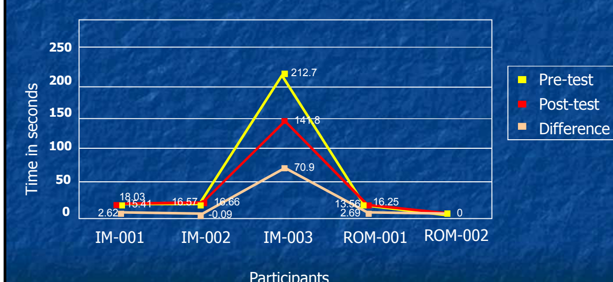
Results

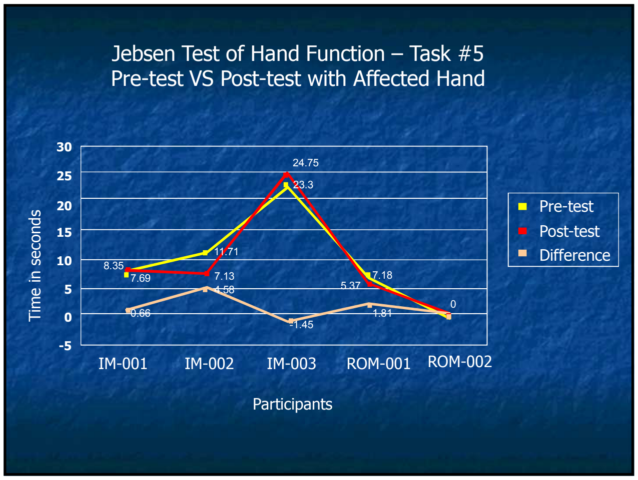
- Study remains an on-going process
- Preliminary findings
  - 3 of 7 Jebsen subtasks (affected hand)
  - 9-Hole Peg Test (both hands)
  - IM Long Form Assessment

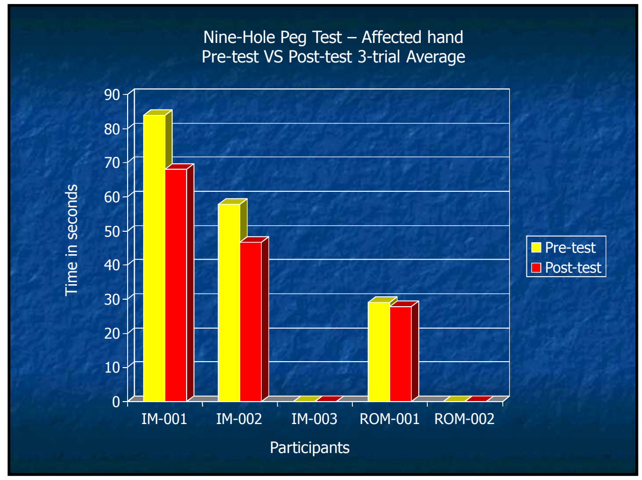
Jebsen Test of Hand Function – Task #1
Pre-test VS Post-test with Affected Hand

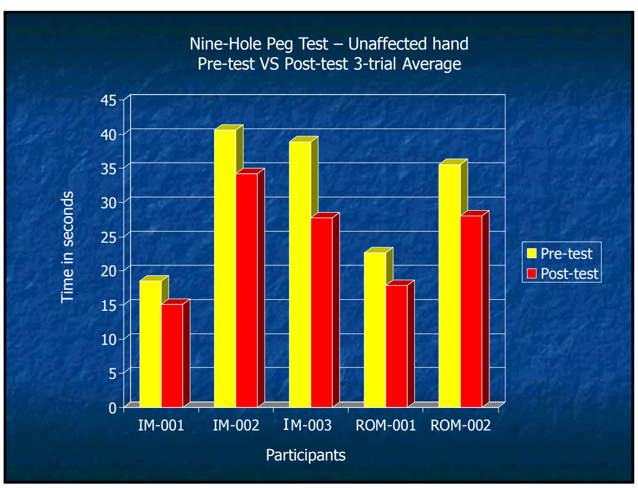


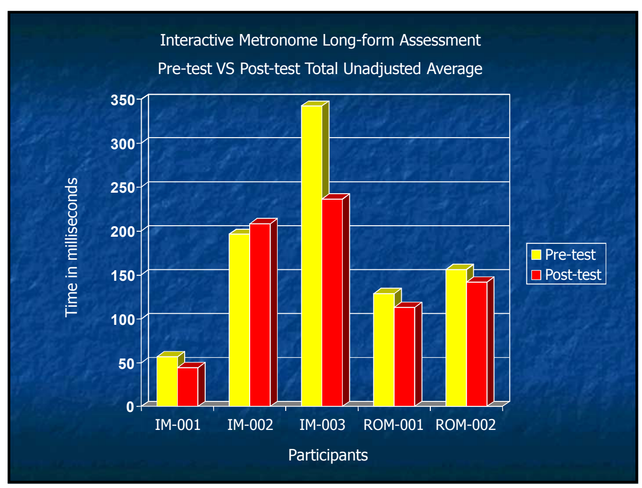
Jebsen Test of Hand Function – Task #4
Pre-test VS Post-test with Affected Hand




Jebsen Test of Hand Function – Task #5
Pre-test VS Post-test with Affected Hand
Time in seconds
30
25
20
15
10
5
0
-5
24.75
23.3
8.35
7.69
11.71
7.13
4.58
0.66
-1.45
5.37
7.18
1.81
0
IM-001
IM-002
IM-003
ROM-001
ROM-002
Participants
Pre-test
Post-test
Difference


Nine-Hole Peg Test – Affected hand
Pre-test VS Post-test 3-trial Average
Time in seconds
90
80
70
60
50
40
30
20
10
0
IM-001
IM-002
IM-003
ROM-001
ROM-002
Participants
Pre-test
Post-test


Nine-Hole Peg Test – Unaffected hand
Pre-test VS Post-test 3-trial Average
Time in seconds
45
40
35
30
25
20
15
10
5
0
IM-001
IM-002
IM-003
ROM-001
ROM-002
Participants
Pre-test
Post-test


Interactive Metronome Long-form Assessment
Pre-test VS Post-test Total Unadjusted Average
Time in milliseconds
350
300
250
200
150
100
50
0
IM-001
IM-002
IM-003
ROM-001
ROM-002
Participants
Pre-test
Post-test

Conclusions

- Preliminary data shows Interactive Metronome's® effectiveness as a treatment modality to improve movement smoothness in individuals post stroke
- Combine IM treatment with standard rehabilitation approaches
- Research is ongoing

Significance of study

- Effectiveness of a treatment approach for individuals with a stroke
- Stroke buckle
- Advances evidence-based practice in occupational therapy
- Local effect & global research prospective

Guerra, 2006; Stube and Jedlicka, 2007

References

- Guerra, A. (2006). *Interactive metronome: A rehab revolution*. Conference presentation presented at 2006 IM Professional conference. Austin, TX.
- Hackel ME, Wolfe GA, Bang SM, CanJeld JS. (1992). Changes in hand function in the aging adult as determined by the Jebsen Test of Hand Function*. Journal of Physical Therapy.* 72*(5)*: 373-377.
- Interactive Metronome (2006). *IM certification provider training manual*. Interactive Metronome, Inc.
- Interactive Metronome (1 January 2008). *IM certification course provider training guide.* IM Certification Course: 19 January 2008; Raleigh, NC.
- Libkuman, T.M., Otani, H. & Steger, N. (2002). Training in timing improves accuracy in golf. *The Journal of General Psychology*, 129*(1)*, 77-96.
- Shaffer, R.J. et al. (2001). Effect of interactive metronome training on children with ADHD. *The American Journal of Occupational Therapy*, 55*(2)*, 155-162.
- Stube, J.E. & Jedlicka, J.S. (2007). The acquisition of integration of evidence-based practice concepts by occupational therapy students. *The American Journal of Occupational Therapy*, 61*(1)*, 53-61.
- Tri-State Stroke Network. (2007). *Facts and Statistics.* Retrieved on 7 Nov 2007 from http://www.tristatestrokenetwork.org/facts.html.

The Journal of General Psychology, 2002, *129*(1), 77–96

Training in Timing Improves Accuracy in Golf

TERRY M. LIBKUMAN
HAJIME OTANI
Department of Psychology
Central Michigan University

NEIL STEGER
Preventive Health, Inc.
Grand Rapids, Michigan

ABSTRACT. In this experiment, the authors investigated the influence of training in timing on performance accuracy in golf. During pre- and posttesting, 40 participants hit golf balls with 4 different clubs in a golf course simulator. The dependent measure was the distance in feet that the ball ended from the target. Between the pre- and posttest, participants in the experimental condition received 10 hr of timing training with an instrument that was designed to train participants to tap their hands and feet in synchrony with target sounds. The participants in the control condition read literature about how to improve their golf swing. The results indicated that the participants in the experimental condition significantly improved their accuracy relative to the participants in the control condition, who did not show any improvement. We concluded that training in timing leads to improvement in accuracy, and that our results have implications for training in golf as well as other complex motor activities.

Key words: golf swing, performance, timing

GOLFERS are constantly looking for ways to improve their performance. One of the ways in which they attempt to accomplish this is through the use of the modern or “high-tech” golf club. Although it is not clear whether performance is enhanced with the modern club, this quick-fix approach is popular, as evidenced by the millions of dollars spent annually on such clubs. The second way of trying to improve performance is through instruction. This approach is also popular, as witnessed by the numerous swing instructors (the so-called swing gurus), schools and academies, magazines, videos, and books devoted to improvement in golf. However, as with the modern golf club, it is not clear what impact instruction has on performance.

Golf aids, commonly used in conjunction with instruction, are another way in which golfers try to enhance performance (Wiren, 1995). There are numerous golf aids on the market. For example, a golfer who believes that he or she has a problem with wrist movement may use an aid (worn on the hand and wrist) that allows only for the appropriate movement. This approach is also popular (witness the common caricature of the golfer weighted down with a multitude of golf aids) but, like the other performance-enhancing approaches, there is little, if any, evidence to support the efficacy of this one.

In contrast to the applied approaches directed toward the improvement of golf performance, there is another approach, in which researchers are more concerned with understanding the nature of the golf swing (e.g., Cochran, 1992, 1995; Cochran & Stobbs, 1968; Hay, 1978; Jorgensen, 1994). This approach implies that understanding the golf swing will lead to its improvement and ultimately to lowered golf scores. Also for researchers, the golf swing, because of its complex nature, poses some interesting intellectual challenges.

Cochran and Stobbs (1968) attempted to simplify the complexity of this phenomenon by modeling the golf swing as a double pendulum system in which two levers rotate about a fixed pivot. The fixed point is between the golfer's shoulders, and it is fixed only in the sense that it does not change planes. The one lever is an upper lever and corresponds to the arms and shoulders swinging around the fixed point. The other lever is a lower lever and corresponds to the movement of the golf club. The two levers are hinged in the middle by the wrists and the hands. A fundamental assumption of this model is that, for the levers to work effectively, it is essential that the levers be timed. In other words, to transfer the maximum amount of energy to the club head at impact, the lower and upper levers must work in synchrony. Therefore, acquisition of this skill, particularly at the expert level (Ericsson, 1996; Ericsson & Lehmann, 1996), requires extensive and effortful practice, not only to learn the basic swing movements but also to time them. Furthermore, we assume that without any additional major changes in the basic movements of the golf swing (for example, changing the golfer's swing plane through training or instruction), the skill must continue to be "fine-tuned" or timed for the golfer to maintain the high level of reliability that is required for successful performance. In fact, a basic assumption made by many professional golfers is that the only practice that should occur immediately before a competitive event is fine-tuning, and that the major downfall in actual competition (with its inherent stresses and pressures) is the failure to maintain proper timing.

This research was funded by Interactive Metronome®, Orlando, FL. Interactive Metronome is a registered trademark of Synaptec, LLC dba Interactive Metronome®, 500 Winderley Place, Suite 108, Maitland, FL 32751. United States Patent #5,529,498 & 5,743,744. Web site: interactivemetronome.com

A portion of the results were presented in Golf Research News. *The authors thank James Cassily, Stanley Greenspan, Sonya Sheffert, Roberto Tuchman, Richard Ferraro, and anonymous reviewers for their helpful comments on the draft of this article.*

Address correspondence to Terry Libkuman, Department of Psychology, Central Michigan University, Mt. Pleasant, MI 48859; libku1tm@cmich.edu (e-mail).

It is important to emphasize that even though there is a scientific body of knowledge about the golf swing, there is little empirical literature concerning the timing properties of the golf swing. This is in direct contrast to the enormous importance that is attached to timing by instructors (e.g., Leadbetter, 1990, 1993) and golfers (e.g., Nicklaus, 1974; Watson, 1998). In fact, it would be a rare event to select any issue of any popular golf magazine (e.g., *Golf* and *Golf Digest*) and not find an article devoted to timing. In the present experiment, therefore, we examined this aspect of the golf swing. In particular, we asked whether extensive training in timing would improve performance accuracy. We chose accuracy over distance as the major dependent measure because even though distance is an important determinant of performance (Cochran & Stobbs, 1968), greens in regulation (an index of accuracy) accounts for more of the variance in golf scores than does any other single measure (Riccio, 1995).

There are at least three indications that training in timing might improve the golf swing performance. Jagacinski, Greenberg, and Liao (1997) found evidence that the age-related decline in golf performance may be explained by the differences in timing, rhythm, and tempo between young and older adults. The researchers referred to timing as those forces that are applied to the golf club during the swing. In contrast, tempo referred to the overall speed of the swing, and rhythm referred to the cycle of speeding up and down of the swing. In the Jagacinski et al. study, young and older adults were asked to swing an eight iron in order to hit a plastic ball that was placed on a rubber tee. The speed and force pattern of the club head was measured by a miniature accelerometer attached to the club head. Jagacinski et al. decomposed the swing by analyzing the force patterns into six phases: (a) beginning of the swing, (b) backswing, (c) downswing up to the maximum force, (d) downswing from the maximum force to impact, (e) impact to the resting level, and (f) the resting level to the maximum force during the follow-through. By measuring the duration of these phases, they were able to test the hypothesis that older adults swing the club too quickly or at too fast a tempo relative to younger adults. Their analyses partially supported the hypothesis: Older adults exhibited a shorter overall shot duration than did younger adults, even though the difference was only marginally significant. Rhythm, measured by the duration of each of the six phases, also showed age differences. The older adults, relative to the younger adults, exhibited shorter intervals during the beginning of the swing, from impact to resting, and from the resting level to the maximum force during the follow-through. Jagacinski et al. interpreted these results as indicating that for younger golfers, the club head reaches its peak maximal force just before impact, whereas for older golfers the club head reaches its peak maximal force earlier in the swing. The obvious implication is that getting the peak maximal force to occur just prior to impact for the older golfers should improve their performance. Interestingly, the amount of force was roughly the same for both groups. Thus, the findings of Jagacinski et al. indicate that timing is important in the golf swing and that age-related declines in golf performance may be due to this factor. On the basis of their results, these authors suggested that training in timing might improve one's golf swing. In particular, they suggested that slowing down the swing and maintaining this same tempo for all shots would be an effective strategy for improving performance.

Another indication that training in timing may improve the golf swing is based on studies that investigated the effects of transcranial stimulation on timing. These studies indicated that by stimulating the motor cortex, a voluntary motor act could be delayed without affecting the intention to act (Day, 1996). Day and colleagues (Day, Dressler, et al., 1989; Day, Rothwell, et al., 1989) administered transcranial stimulation in two ways. One was a short-duration, high-voltage electrical stimulus that passed through an electrode attached to the scalp; the other stimulus was a pulsed magnetic field delivered through a flat, circular coil held on the head. The stimulation was delivered 100 ms after the onset of a "go" signal. The results showed that both types of stimulation delayed the onset (approximately 50 ms) of the motor movement (i.e., wrist flex and wrist extension). Furthermore, the electromyographical pattern of agonist/antagonist muscle activation (i.e., contracting muscles that are resisted by other muscles) was similar between trials with or without the stimulation. The latter observation indicated that the stimulation did not affect the way in which their voluntary movement was produced. In contrast, stimulation to the peripheral nerve produced different results. When the median nerve at the elbow was stimulated, there was no delay in the onset of muscle activity. The stimulation suppressed only the first burst of agonist muscle activity. On the basis of these observations, Day and colleagues (Day, Dressler, et al., 1989; Day, Rothwell, et al. (1989) concluded that stimulation per se does not cause the delay. Day, Rothwell, et al. (1989) also asked whether the stimulation delays the onset of movement by delaying one's intention to act. To test this hypothesis, they instructed participants to flex both wrists while receiving stimulation to the motor cortex from only one side of the brain. The rationale for this treatment was that if the stimulation delays the participants' intention to act, then the unilateral stimulation would delay the activation of the muscles for both wrists. On the other hand, if the stimulation delays the movement by affecting an executive process that controls the nerve pathways, the unilateral stimulation would delay only the movement of the limb contralateral to the stimulation. The results showed that the delay of movement was greater for the contralateral limb than for the ipsolateral limb. They concluded that the cortical stimulation does not affect one's intention to act. Instead, stimulation delays movement by affecting the executive process that sends signals to the muscle.

Day (1996) interpreted these results to mean that transcranial stimulation inhibits the motor cortex to initiate the movement. However, this does not explain the result that the normal movement returned after the cortical inhibition was over. To explain this, Day proposed a hierarchical model of timing consisting of two partially independent components. One is a high-level process that prepares the movement and instructs the motor cortex to release the movement. The second is a subordinate level process that refines the precise timing of the movement. It is the second process that determines when the instructions to move relevant muscles would be sent. According to Day, the important property of this model is that "our limbs would not necessarily move when we tell them" (p. 233). For our purpose, this implies that practice may be needed to refine the coordination between one's intention to act and the precise timing of the act itself.

A more recent view of sensory and motor timing also proposes a common neural mechanism to represent temporal properties of perceived events and motor movements (Meegan, Aslin, & Jacobs, 2000). Research has suggested that the cerebellum may play an important role in representing sensory and motor timing (Ivry & Keele, 1989; Jueptner et al., 1995). In support of this view, Meegan et al. showed that motor timing could be improved by sensory timing training. In that study, participants were asked to use their right thumbs to press a button twice in succession with a prespecified interpress interval. The sensory training consisted of discriminating between a short and long interval between two tones. The researchers found that even though the sensory training did not involve motor movements, motor performance improved significantly after the training. On the basis of these results, Meegan et al. concluded that sensory timing training alters motor timing because a common neural mechanism is used to represent timing for the sensory and motor systems.

On the basis of the considerations mentioned in our literature review, we thought that it would be useful to examine the notion that extensive training in timing would improve performance in golf. The design of the present study was relatively simple. First, all participants were pretested, with accuracy as the measure of golf performance. Second, the participants were assigned to the experimental or control condition. The experimental-condition group received approximately 10 hr of training with a specialized metronome (Interactive Metronome®). The Interactive Metronome®, unlike other metronomes, uses auditory feedback to train an individual to match a variety of movements to a steady beat. The control-condition group read golf instruction literature. Third, after 5 weeks, both groups were posttested with the same procedure and measure that were used in the pretest. We hypothesized that training in timing would improve accuracy.

The more important consideration in the design of the study was the timing parameter. What value should be selected? Furthermore, should the value remain constant or should it vary across training? Because there are no known empirical studies that have tested for the effects of timing on golf, and little, if any, theoretical guidance, we had to set the timing parameter largely on the basis of experience and intuition. In agreement with the suggestion of Jagacinski et al. (1997), we fixed the value at a relatively slow pace of 54 beats per minute (bpm) for all of the motor tasks across all of the training sessions. We assumed performance problems associated with the timing of the golf swing were largely due to tempo, and that extensive training at the slow pace of 54 bpm would improve tempo. Finally, we did not ask participants to practice with a golf club because we assumed that movements are stored in the central nervous system as general motor programs (e.g., Schmidt, 1975), and therefore the training does not have to be task specific. A recent study by Meegan et al. (2000) also supports the assumption that training in timing does not require motor movements.

Method

Participants

We recruited participants via advertisements that were posted in local golf retail shops, at driving ranges, and in the pro shops of area country clubs. The advertisements stated that participants were needed for a golf training technology study and that the study was designed to evaluate the effectiveness of a golf skills training aid on golf shot accuracy. Participants were informed of the schedule and time requirements of the study. To qualify for participation, interested individuals had to be 25 years of age or older and had to possess at least a basic skill level in golf. The first 50 individuals who met these requirements were selected and randomly assigned to the 2 conditions with the restriction that each condition contained 25 participants. Of the 50 participants who started the study, 9 did not complete it. Further, 1 participant from the experimental group was randomly excluded to equalize the numbers of participants in the experimental and control conditions. The final sample therefore consisted of 6 women and 34 men who ranged in age between 25 and 61 years ($M = 37$, $SD = 11.57$). Unfortunately, the random assignment produced a significant age difference, $t(38) = 4.34$, $p < .001$, between the experimental ($M = 45$, $SD = 11.62$) and control groups ($M = 31$, $SD = 6.43$). (One participant in the control condition did not report her age.) To statistically control for this variable, we analyzed the data using age as a covariate. Participants were informed that, if they completed the study, they would receive a gift certificate for golf equipment or clothing and that they would be competing for two $100 bonus prizes. Finally, participants were informed as to the risks and benefits of participation before they signed informed consent.

Apparatus

Pre- and posttest accuracy was measured using a Full Swing Golf Simulator™ located in an indoor 10 ft × 10 ft × 20 ft booth in a local retail golf shop. The indoor booth allowed for a controlled testing environment. As the name implies, the Full Swing Golf Simulator™ allows the golfer to execute a full swing and to hit a golf ball onto a screen that contains a picture of a golf hole including the tee box, fairway, and green with a pin and flag. The golfer can play a simulated round of golf at a number of famous golf courses. The simulator estimates the distance and direction for each shot and records the score for each hole. The simulator also provides for each shot a visual ball path trajectory line or a visual image of the flight of the golf ball from impact until the ball is stationary. Particularly important for the present study, the Full Swing Golf Simulator^{tm} contains a duel-tracking system that cycles more than 2 million infra-red beams per second. As a consequence, the simulator is able to accurately monitor ball flight within 0.1 in. The measure of accuracy used in the present study for each golf shot was the distance in feet between the golf ball and the pin. Finally, the simulator requires that the approximate box-to-pin yardage be estimated and preset for each club. For example, a golfer hitting a nine iron would estimate and set his distance at 125 yards, a five iron at 170 yards, and so forth.

The Interactive Metronome® was used to train and analyze the golfer's ability to match a variety of movements to a steady beat. The Interactive Metronome® is a computer program for Windows 95/98 with peripherals, which include standard stereo headphones and a set of motion-sensing triggers. The trigger set plugs into the computer's serial port and includes a hand glove and a footpad. One trigger is attached to the participant's hand with a Velcro^{tm} strap. When the participant claps or pats a hand, the attached trigger sends a signal to the program. A second trigger is contained in a floor pad on which the participant steps or taps. The computer program produces an auditory fixed reference beat. The beat can be set at any number of beats per minute. Participants are required to complete various hand and foot exercises in synchrony with the beat. The objective on the part of the participant is to move his or her limb at the same time as that set on the metronome. In other words, the participant attempts to pat or tap his or her hand or foot at the exact moment of the beat.

The program immediately analyzes the timing relationship between the participant's movements and the beat to the nearest millisecond. The tone of the beat (C6) is in monophonic and thus is spatially perceived as occurring in the center of the headphones. Movements include variations of clapping hands together, tapping the right or left hand on the side of the leg, tapping both toes or heels on the footpad, or tapping the right or left toe or heel on the footpad. The program produces different discriminative sounds that are based on the pitch and placement in the headphones. These reference pitches are tailored to guide the participant. The program transposes the timing information of each movement into one of the recognizable sounds. Each sound is a representation of when the movement occurred in relation to the beat. An early movement (i.e., a movement that precedes the beat) generates a low pitch tone in the user's left ear. A late movement (i.e., a movement that follows the beat) generates a higher pitch tone in the right ear. A movement that matches the beat within ±15 ms generates a higher pitched tone in the center of the headphones and is simultaneously perceived in both ears. A participant's timing score is the difference in milliseconds between the moment the beat sounds and the participant's tap.

All of the experimental-condition participants received their training in a room that contained five desktop computers arranged at the points of a pentagon. The computers, monitors, keyboards, and other materials were placed on tables, each with a chair. There were no partitions between the stations. The spacing and arrangement of the stations allowed the participants to stare ahead and not see anyone else working. The participants were also not likely to be disturbed by extraneous sounds because they were wearing headphones.

Procedure

The participants were randomly assigned to the two conditions prior to the pretest. The pretest was completed for all participants on two consecutive Saturdays in the month of May. Each participant was scheduled for a 1-hr appointment on one of the Saturdays at his or her convenience. Participants were informed that the pretest would take about 1 hr and that they should bring their own golf clubs. They were also informed that the type of clothing and shoes worn during the pretest should be worn during the posttest. The participants played the same hole under the same conditions (Troon North Course, AZ, Hole #1) for all shots using the same balls, driving mats, and rubber tees. The pretest consisted of 15 shots each with their nine, seven, and five irons, and the driver for a total of 60 shots. There was a 1-min rest period between each set of 15 shots. The participants were permitted to go through their normal warm-up routine and take as many as 10 shots before beginning the pretest.

In the actual pretest, participants began by setting the distance from the tee box to the pin. The experimenters informed the participants that the selected distance for each club would also be used for the posttest and that they would be required to use the same club. The participants were then instructed to aim for the pin and to proceed at their own pace. The experimenter recorded each score (i.e., the distance in feet from the pin). Finally, all participants were informed that they were strictly prohibited from practicing with any of the clubs that were used during the pretest as well as receiving any instruction or lessons during the study.

The participants in the experimental-condition group (n = 20) received 10 hr of Interactive Metronome® training in 12 sessions of 50-min each. The sessions began the day after the completion of the pretest. They were scheduled throughout the day and early evening for the next 5 weeks. The schedule included weekdays and weekends. Participants scheduled the sessions at their convenience with the stipulation that they could not complete more than 1 training session per day, and that they needed to complete the entire training sequence by the end of the 5-week period. All of the experimental participants were trained in the same room that contained the five computer stations. An experimenter was present for all sessions. Up to 5 participants could be trained simultaneously with one experimenter monitoring their activities by sitting on a bar stool that was placed in the middle of the pentagon. Six experimenters (including the experimenter who collected the pre- and posttest data) were paid and trained in the use of the Interactive Metronome.® All of these experimenters had completed the actual training themselves. There was no attempt to balance experimenters with participants or training sessions. The experimenters simply signed up for scheduled times that were convenient for them and compatible with participant times. The primary duties of the experimenters were to greet the participants and ensure that they signed in and selected the correct daily training schedule. Experimenters also monitored and corrected, if necessary, any technical problems with the equipment, recorded data that were not recorded by the software program, and made sure that the participant scheduled his or her next training session before leaving. Finally, during the first session, experimenters modeled the use of the equipment and the proper technique for each of the prescribed motor movements that were later required of the participants.

Before each training session began, participants were required to sign in and select the appropriate training schedule for the day and to enter some demographic information (e.g., name, age, sex) into the computer. The experimenter attached the hand sensor to the participant's hand, and placed the headphones properly on the head. The experimenter stressed the importance of using controlled, smooth (nonballistic) motions in matching the movement to the steady reference beat. The experimenter also emphasized that participants should not aim, think about, adjust their motions, or listen to the guidance sounds, but rather focus their attention on the steady beat, and whenever they got off beat to refocus their attention on the beat. These instructions were also posted beside each computer station.

The beat of the metronome was set at 54 bpm for all 12 sessions. For each of the tasks within each of the 12 training sessions, concurrent, temporally based, guide sounds continually indicated that the participant was on target, early, or late. At the beginning of the first session and at the end of the last training session, participants were administered 30- to 60-s tests on each of the 13 movements that were used in the training sessions. Guidance sounds were not used during the testing, with the exception of one additional task (the 14th), which was a repeat of clapping both hands with the standard guide sounds. The test took about 10 min to complete. Two dependent measures were recorded for each task: One was the mean number of milliseconds across the 14 tasks the participant deviated from on-target performance, and the other was the highest number of times in-a-row (IARs) that the participant was able to stay within ±15 ms of the reference beat.

Before beginning the 10-min test, the experimenter placed the hand sensor on the participant's hand, and the foot sensor was placed on the floor. Then the experimenter modeled the appropriate movements. There were no exercises that paralleled the motions in the golf swing.

The first 4 tasks in the 10-min test involved the hands. In the 1st task (clapping hands), participants were instructed to make circles of about 10-in. in diameter with the hands coming together on the beat and to continue the circular path without stopping after the beat. The 2nd task was identical to the 1st with the exception that the early, late, and on-target guidance sounds were presented. The guidance sounds were presented only for the 2nd task in the 10-min test. The 3rd

TABLE 1
Training Schedule

Task	Session 1	2	3	4	5
Clapping hands	180 (1)	385 (1)	500 (1)	1000 (1)	1000 (3)
Preferred hand	180 (2)	385 (2)	500 (2)		
Nonpreferred hand	180 (3)	385 (3)		500 (3)	
Both toes	180 (4)	385 (4)	500 (3)	500 (2)	
Preferred toe	180 (5)	385 (5)			
Nonpreferred toe	180 (6)	385 (6)			
Both heels		385 (7)			
Preferred heel					
Nonpreferred heel[a]					
Preferred hand and nonpreferred toe			250 (4)		500 (1)
Nonpreferred hand and preferred toe			250 (5)		500 (2)
Choice			500[b] (6)	250[b] (4)	
Free style					500[c] (3)
Total beats	1080	2695	2500	2250	2500

Note. Value in each cell indicates the prescribed total number of beats that were to be completed. Value in parentheses indicates the order in which the task was presented. [a]The use of the nonpreferred heel occurred only in the both heels and free style tasks. [b]Participant could choose any task that had been previously performed. [c]Participant was required to start with clapping hands, move to preferred hand, then preferred toe, nonpreferred toe, and end with both toes, all within 500 beats. [d]Participant was required to complete three sequences: Sequence 1, 4 beats clapping hands alternating with 4 beats preferred hand for 250 beats; Sequence 2, 4 beats clapping hands alternating with 4 beats both toes for 250 beats; and Sequence 3, 2 beats clapping hands alternating with 2 beats both toes for 500 beats. [e]Participant was required to alternate between 8 beats clapping both hands and 8 beats both toes. [f]Participant could switch between any of the tasks with the restriction to limit switching to every 100 beats.

and 4th tasks involved using either the preferred or nonpreferred hand and required that the participant, using the same circular motion, tap his or her hand on his or her leg on beat. The next 3 tasks involved the toes. In the 5th task, participants were instructed to face the floor trigger with both toes about 2 to 3 in. away from the trigger. They were instructed to start by lifting one foot and tapping that toe on the trigger with the beat and to return that foot to the previous position between beats, then tap the other toe on the next beat, and so forth. Tasks 6 and 7 involved the same movement but with only the preferred or nonpreferred toe, respectively. The next 3 tasks involved the heels. In the 8th task, participants were instructed to face away from the floor trigger with both heels about 2 to 3 in. away from the trigger and to start by lifting one foot and tapping that heel on the trigger on the beat, and return that foot to the previous position between beats, and then tap the other heel on the next beat, and so forth. Tasks 9 and 10 involved

Session 6	7	8	9	10	11	12	Σ
1000 (1)		1500 (1)		2000 (1)			7565
	1000 (1)						2065
		500 (2)					1565
1000 (3)				500 (2)			3065
	500 (2)				250 (2)		1315
					250 (3)		815
500 (2)							885
		500 (3)					500
							000
			1000 (1)				1750
			1000 (2)				1750
							750
	1000[d] (3)		500[e] (3)		2000[f] (1)	2000[f] (1)	6000
2500	2500	2500	2500	2500	2500	2000	

the same movement but with only the preferred or nonpreferred heel, respectively. The next 2 tasks involved combinations of movements. In Task 11, the preferred hand and nonpreferred toe were combined. Participants were instructed to face the floor trigger, tap their preferred hand against their leg on one beat, then tap the toe of the opposite (nonpreferred) foot on the floor trigger on the next beat and then to continue to alternate. In Task 12, the nonpreferred hand and preferred toe were combined with the same movements outlined in Task 11. In the final 2 tasks, balancing was added. In Task 13, participants were required to balance on their preferred leg while tapping the toe of their other foot on the floor trigger on each beat, and in Task 14, they had to switch to the nonpreferred leg.

After the completion of the 10-min test, the training sessions began. The purpose of the training was to increase the timing accuracy. Table 1 provides the training schedule. The development of this training schedule was based on three assumptions. First, we incorporated variability in the tasks that were required because we thought it would be more likely to generalize or transfer to another motor activity (Schmidt, 1988), in this case the golf swing. In other words, participants would become more sensitive to the timing properties necessary to execute this motor response. Second, although the total number of beats was relatively consistent across sessions (the number of beats required for testing are not included in Table 1), we increased the number of beats per task and decreased the number of tasks across sessions, assuming that this type of extended training on

TABLE 2
Mean IAR and Deviation From the Target in ms as a Function of Task and Test, Pretest and Posttest

	IAR		Target deviation	
Task	Pre	Post	Pre	Post
1. Both hands				
M	3.40*	5.85*	48.86*	21.17*
SD	2.30	2.68	19.32	6.86
2. Both hands with sounds				
M	2.70*	6.30*	71.54*	19.85*
SD	1.81	1.94	43.70	6.30
3. Preferred hand				
M	2.75*	4.50*	42.12*	25.19*
SD	1.83	2.59	17.00	13.02
4. Nonpreferred hand				
M	2.50*	4.00*	41.67*	22.65*
SD	1.67	2.34	19.45	7.57
5. Both toes				
M	1.90*	3.60*	68.99*	26.40*
SD	1.25	1.47	48.26	10.67
6. Preferred toe				
M	2.05*	3.70*	53.24*	27.38*
SD	1.19	2.68	24.89	9.82
7. Nonpreferred toe				
M	2.00*	3.90*	58.32*	27.08*
SD	1.26	2.05	32.67	10.25
8. Both heels				
M	1.65*	2.90*	71.43*	32.17*
SD	1.09	1.83	36.02	16.66

(table continues)

a single task would lead to an increase in the ability to maintain focus on the task as well as when executing the golf swing. Third, because of the positive relationship between the amount of practice and skilled performance (Ericsson, 1996; Schmidt, 1988), we assumed that by providing 10 hr of training (a total of 28,025 beats plus the beats during testing), the training in timing would be more likely to transfer to the golf swing. Finally, because of the repetitive nature of the training, participants in the experimental group were provided with motivating instructions beginning with the 3rd session and ending with the 11th session. These instructions urged them to decrease their millisecond average and increase their IARs. Furthermore, participants were informed that their millisecond averages and IARs would be ranked and posted and that the top two performing individuals would receive a $100 gift certificate for golf equipment or clothing.

TABLE 2 (Continued)

Task	IAR Pre	IAR Post	Target deviation Pre	Target deviation Post
9. Preferred heal				
M	1.60*	2.85*	96.74*	36.08*
SD	1.43	1.50	99.04	15.62
10. Nonpreferred heel				
M	1.55	2.30	76.07*	38.37*
SD	1.32	1.49	49.77	18.52
11. Preferred hand and nonpreferred toe				
M	1.15*	2.15*	97.75*	42.14*
SD	0.75	0.93	41.76	16.24
12. Nonpreferred hand and nonpreferred toe				
M	1.15*	2.50*	100.10*	34.69*
SD	0.88	1.28	57.17	11.04
13. Balance with preferred foot and tap with nonpreferred toe				
M	1.65	2.75	70.63*	33.08*
SD	1.27	2.00	38.60	14.89
14. Balance with nonpreferred foot and tap with preferred toe				
M	1.55*	3.15*	61.64*	25.15*
SD	0.94	1.53	28.19	6.77

Note. IAR = number of items in-a-row.
*$p < .05$.

In contrast to the participants in the experimental group, the participants in the control group received a letter indicating that the attached 12 pages of golf tips were to be read at least once a day before the posttest. The golf tips were taken from popular golf magazines and books and were authored by prominent professional golfers and instructors. The participants were also informed that after completing the posttest they would receive a golf certificate. The control participants were not contacted again until they were scheduled for the posttest.

Results

Unless otherwise specified, the significance level was set at .05 for all of the analyses. We first determined whether the participants in the experimental group made a significant improvement in timing. Table 2 shows how participants performed on the tasks in the 10-min test before and after the training. As mentioned earlier, IARs and the milliseconds from the target were used to index the participants' timing. The table indicates that for both measures, participants performed better on the posttest than on the pretest (see Table 2). A 2 (test: pretest and posttest) × 14 (task: 14 different tasks on the 10-min test completed) repeated-measures analysis of variance (ANOVA) on the IAR scores indicated that the effects of test, $F(1, 19) = 145.61$, $MSE = 2.56$, task, $F(13, 247) = 12.46$, $MSE = 2.80$, and the Test × Task interaction, $F(13, 247) = 2.40$, $MSE = 2.08$, were significant. The interaction simply indicated that the amount of improvement differed across tasks. A priori *t* tests performed on each task indicated that all tasks except for 2, tapping with the nonpreferred heel and tapping with the nonpreferred toe while balancing on the preferred foot, showed significant improvement from pre- to posttest. However, improvement was marginally significant for these 2 tasks ($p < .10$). Similar results were obtained with the deviation from the target measure. A 2 (test: pretest and posttest) × 14 (task: 14 different tasks on the 10-min test) repeated-measures ANOVA revealed that the effects of test, $F(1, 19) = 53.16$, $MSE = 4031.07$, task, $F(13, 247) = 9.48$, $MSE = 661.68$, and the Test × Task interaction, $F(13, 247) = 3.27$, $MSE = 675.39$, were significant. A priori *t* tests showed that all 14 tasks showed significant improvement from pre- to posttest. In summary, both measures indicated that the metronome training improved participants' timing.

Next, we analyzed the accuracy scores. We measured accuracy by the distance in feet between the pin and the ball's final resting place. The scores were averaged over 15 trials for each club for each participant. Table 3 displays the mean accuracy as a function of club, treatment group, and test (pre and post). As Table 3 indicates, the overall performance of the experimental group was better than that of the control group. Also, the accuracy differed between clubs. Figure 1 further shows the mean improvement that occurred between pre- and posttest as a function of club and treatment group. As shown, performance improved for the experimental condition for all clubs. In contrast, little or no improvement occurred for the control condition. These observations were confirmed by a 2 (group: experimental and control) × 4 (club: nine iron, seven iron, five iron, and driver) × 2 (test: pre- and posttest) mixed-design ANOVA. The results revealed that the main effect of club, $F(3, 114) = 106.14$, $MSE = 2323.48$, and the Group × Test interaction, $F(1, 114) = 4.42$, $MSE = 2598.87$, were significant. The main effect of group, $F(1, 39) = 3.10$, $MSE = 17308.23$, and test, $F(1, 114) = 3.13$, $MSE = 2598.87$, approached significance ($p < .10$). A priori independent *t* tests indicated that the treatment groups did not differ from each other on the pretest, $t(39) < 1$. However, on the posttest, the experimental group was significantly more accurate than the control group, $t(38) = 2.97$. Furthermore, paired-sample *t* tests indicated that there was a significant increase in accuracy between the pre- and posttest for the experimental group, $t(19) = 2.69$. No improvement occurred in the control group, $t(19) < 1$.

Because there was a significant difference in age between the experimental and control groups, we conducted another analysis on accuracy using age as a covariate. We also used the mean estimated distance across four clubs as a covari-

TABLE 3
Pretest and Posttest Mean Accuracy in Feet as a Function of Group, Club, and Testing

	Club									
	9 Iron		7 Iron		5 Iron		Driver		Overall	
Group	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Experimental										
M	66.32	57.00	76.95	58.31	114.45	82.14	186.86	158.87	111.15	89.08
SD	39.78	24.49	55.21	22.29	80.60	35.64	113.02	96.60	66.99	39.63
Control										
M	78.64	78.20	88.01	93.45	122.28	126.75	212.26	210.43	125.32	127.21
SD	44.81	32.97	43.88	38.77	62.86	66.61	70.52	66.20	48.77	41.59

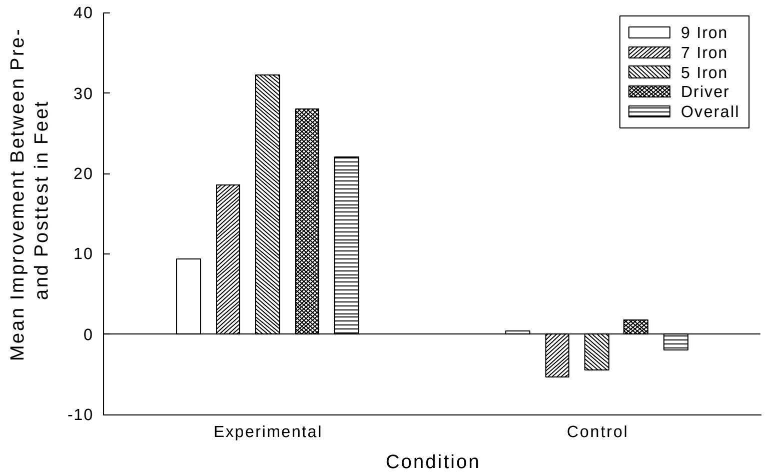


FIGURE 1. Mean improvement as a function of club and treatment group.

ate. As mentioned earlier, each participant determined at the pretest how far he or she would be able to hit the ball with each club. We expected the estimated distance to reflect each participant's expertise with playing golf. By using this variable as a covariate, we attempted to equate the level of expertise between the experimental and control groups. A 2 (group: experimental and control) × 4 (club: nine iron, seven iron, five iron, and driver) × 2 (test: pre- and posttest) mixed-design analysis of covariance (ANCOVA) indicated that the effects of club, $F(3, 105) = 6.35$, $MSE = 2318.47$, test, $F(1, 35) = 5.07$, $MSE = 2462.79$, and the Group × Test interaction, $F(1, 105) = 4.72$, $MSE = 2462.79$, were significant. Further analyses indicated that these results were similar to the previous accuracy analysis.

We also correlated age with IAR and millisecond deviation scores to rule out further that age was a factor in producing improvement. We computed a correlation between age and improvement that occurred in IAR and millisecond deviation scores between pre- and posttest (pre–post) on each task. None of the correlations except one was significantly different from zero. The only significant correlation occurred with the millisecond deviation score on the task that required tapping with the nonpreferred toe, $r = .55$. The positive correlation indicated that improvement was greater for older adults relative to younger adults. However, no other correlations reached significance, indicating that age was an unlikely source of improvement in overall timing. In summary, the results of this study indicate that the training in timing improved accuracy relative to a control group, which did not show any improvement.

Discussion

The results of the present experiment suggest that training in timing improves accuracy in golf. Furthermore, the improvement in performance was consistent across golf clubs. Why does training in timing on an activity that does not mimic the golf swing enhance accuracy in this activity? There are several possibilities.

One obvious answer is that the training improved the golf swing by fine-tuning the timing properties (i.e., tempo and rhythm) of the golf swing. As mentioned in the introduction, the golfing community has attached considerable importance to the notion that timing is an essential property in a successful golf swing. Unfortunately, in the present study, we can only speculate about which timing properties were changed because these properties were not measured. However, we specifically suggest that the training in timing leads to changes in tempo. In support of this notion, Jagacinski et al. (1997) demonstrated that older individuals have faster tempos than younger individuals. These authors also reported that the maximal force of the club head occurs earlier with an older adult than with a younger adult. Note that the mean age of our experimental participants (M = 44) falls somewhat in between the age range (mean ages were not provided) of older participants (60 to 69) and the younger participants (19 to 25) in the study of Jagacinski et al. Thus, it is possible that training improved the tempo of the golfers in our study.

The second possibility is that the training made the coordination between participant's intention and voluntary movement more precise. On the basis of the model of Day (1996), intention to act and voluntary movement are organized in a hierarchical fashion. As Day indicated, the important implication of this model is that our limbs may not move when we intend to move them. It is possible that even without external interference (e.g., transcranial stimulation), the coordination between the motor planning component and the timing component is not perfect. Therefore, fine-tuning between these components may be necessary to produce motor movements that require precise timing. Similarly, it is conceivable that sensory training using the Interactive Metronome® may have modified the temporal representation used for both sensory and motor systems. In support of this hypothesis, our results are consistent with the results of Meegan et al. (2000), which indicate that motor movements are not necessary to improve the temporal properties of the motor movements.

The third possibility is that the improvement was simply an artifact of demand characteristics. Participants in the control group were not asked to come to the laboratory to engage in activities that could possibly improve their golf swing. It is difficult to rule out this possibility without further investigations in which other groups would be tested using other motor exercises. However, we are inclined to believe that the improvement in accuracy had something to do with timing. It is a commonly reported experience that improvement in golf, as in any highly skilled behavior, requires extensive and effortful practice with feedback

(Ericsson, 1996). We therefore doubt that the transient nature of demand characteristics can account for our results. Furthermore, it is important to note that although participants in the control group were provided with golf tips, these participants failed to show any improvement.

In the present study, we provided extensive training by varying the total number of beats across a variety of tasks while maintaining the same number of beats per minute. In future studies, it would be interesting and important to examine the effectiveness of various schedules that include different tasks, durations, and beats per minute. These studies could provide data concerning the most optimal relationship between timing and golf performance. Within this context, it would also be important to include other measures of golf performance, for example, distance in driving and accuracy in putting. Furthermore, future studies should examine the relationship between timing and golf performance by directly measuring some of the temporal properties of the golf swing itself, something that was not done in the present study. Even more ideally, at an individual differences level, it may be possible to determine the number of beats per minute that is most effective in producing the tempo that leads to the most effective performance. In other words, effective performance may depend on temporal properties that are unique to each individual, and the training may need to be tailored to each individual.

Future studies could also take advantage of the golf simulator to separate the distance and direction of the shots. It is possible that training in timing would improve both. Furthermore, the golf simulator is capable of simulating both fairway and green shots. Perhaps timing is more important for one type of shot than it is for the other. Also, in our study, participants were told to ignore feedback (i.e., the guidance sounds) when they were trained with the Interactive Metronome.® It would be interesting to examine whether focusing on feedback would influence the effectiveness of the training.[1]

Finally, the present results provide some interesting implications for other motor activities. If training in timing improves performance by fine-tuning the timing components of a motor movement, then this type of training may be used to improve performance in other activities that require precise timing. Thus, it would be interesting to examine whether Interactive Metronome® training would improve movements in other sports (e.g., basketball, baseball, and tennis) as well as in other endeavors such as flying and typing.

In summary, the results of the present experiment indicated that training in timing improved accuracy in golf. Future research will be necessary for further delineation of the phenomenon and for development of a theory that can explain how the property of timing influences this complex motor activity. However, it is important to note that this is the first experimental demonstration of the effec-

---

[1]We thank an anonymous reviewer for suggesting the future studies mentioned in this paragraph.

tiveness of training in timing on a complex motor activity, and that now there is evidence to indicate that training in timing may improve one's performance in golf. We envision that an instrument such as the Interactive Metronome® could be used not only for overall training in timing but also for fine-tuning one's swing before and during competition. Finally, we agree with Cochran and Stobbs (1968) that the terminology and concepts describing the temporal properties of the golf swing are elusive even though there is nothing more obvious than the gracefulness of a well-timed golf swing.

REFERENCES

Cochran, A. J. (Ed.). (1992). *Science and golf*. London: E & FN Spon.

Cochran, A. J. (Ed.). (1995). *Golf: The scientific way*. Hertfordshire, UK: Aston Publishing Group.

Cochran, A. J., & Stobbs, J. (1968). *The search for the perfect swing*. Grass Valley, CA: The Booklegger.

Day, B. L. (1996). A hierarchical model of motor timing. In M. A. Pastor & J. Artieda (Eds.), *Time, internal clocks and movement* (pp. 223–234). New York: Elsevier.

Day, B. L., Dressler, D., Maertens de Noordhout, A., Marsden, C. D., Nakashima, K., Rothwell, J. C., & Thompson, P. D. (1989). Electric and magnetic stimulation of human motor cortex: Surface EMG and single motor unit responses. *Journal of Physiology, 412*, 449–473.

Day, B. L., Rothwell, J. C., Thompson, P. D., Maertens de Noordhout, A., Nakashima, K., Shannon, K., & Marsden, C. D. (1989). Delay in the execution of voluntary movement by electrical or magnetic brain stimulation in intact man. *Brain, 112*, 649–663.

Ericsson, K. A. (Ed.). (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games.* Mahwah, NJ: Erlbaum.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology, 47*, 273–305.

Full Swing Golf Simulator [Equipment]. (1995). San Diego, CA: Full Swing Golf, Inc.

Hay, J. G. (1978). *The biomechanics of sports techniques* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Interactive Metronome® [Equipment]. (1996). Orlando, FL: Interactive Metronome.®

Ivry, R. B., & Keele, S. W. (1989). Timing function of the cerebellum. *Journal of Cognitive Neuroscience, 1*, 136–152.

Jagacinski, R. J., Greenberg, N., & Liao, M. (1997). Tempo, rhythm, and aging in golf. *Journal of Motor Behavior, 29*, 159–173.

Jorgensen, T. P. (1994). *The physics of golf*. New York: American Institute of Physics.

Jueptner, M., Rijntjes, M., Weiller, C., Faiss, J. H., Timmann, D., Mueller, S. P., & Diener, H. C. (1995). Localization of a cerebellar timing process using PET. *Neurology, 45*, 1540–1545.

Leadbetter, D. (1990). *The golf swing*. New York: Penguin Group.

Leadbetter, D. (1993). *Faults and fixes*. New York: HarperCollins.

Meegan, D. V., Aslin, R. N., & Jacobs, R. A. (2000). Motor timing learned without motor training. *Nature Neuroscience, 3*, 860–862.

Nicklaus, J. (1974). *Golf my way*. New York: Simon & Schuster.

Riccio, L. (1995). What it takes to break 80. In A. Cochran (Ed.), *Golf: The scientific way* (pp. 234–237). Hertfordshire, UK: Aston Publishing Group.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225–260.

Schmidt, R. A. (1988). *Motor control and learning: A behavioral emphasis* (2nd ed.). Champaign, IL: Human Kinetics.
Watson, T. (1998, June). Watson's shortcuts: How to make solid contact. *Golf Digest, 49*, 56.
Wiren, G. (1995). Enhancing performance through golf learning aids. In A. Cochran (Ed.), *Golf: The scientific way* (pp. 205–207). Hertfordshire, UK: Aston Publishing Group.

Manuscript received June 5, 2000
Revision accepted for publication March 6, 2001

Effect of Interactive Metronome® Training on Children With ADHD

Robert J. Shaffer, Lee E. Jacokes, James F. Cassily, Stanley I. Greenspan, Robert F. Tuchman, Paul J. Stemmer, Jr.

Key Words: attention deficit disorder with hyperactivity • coordination training • motor control

Objective. *The purpose of this study was to determine the effects of a specific intervention, the Interactive Metronome®, on selected aspects of motor and cognitive skills in a group of children with attention deficit hyperactivity disorder (ADHD).*

Method. *The study included 56 boys who were 6 years to 12 years of age and diagnosed before they entered the study as having ADHD. The participants were pretested and randomly assigned to one of three matched groups. A group of 19 participants receiving 15 hr of Interactive Metronome training exercises were compared with a group receiving no intervention and a group receiving training on selected computer video games.*

Results. *A significant pattern of improvement across 53 of 58 variables favoring the Interactive Metronome treatment was found. Additionally, several significant differences were found among the treatment groups and between pretreatment and posttreatment factors on performance in areas of attention, motor control, language processing, reading, and parental reports of improvements in regulation of aggressive behavior.*

Conclusion. *The Interactive Metronome training appears to facilitate a number of capacities, including attention, motor control, and selected academic skills, in boys with ADHD.*

Shaffer, R. J., Jacokes, L. E., Cassily, J. F., Greenspan, S. I., Tuchman, R. F., & Stemmer, P. J., Jr. (2001). Effect of Interactive Metronome® training on children with ADHD. *American Journal of Occupational Therapy, 55,* 155–162.

Robert J. Shaffer, PhD, is Adjunct Assistant Professor of Pediatrics and Human Development, College of Human Medicine, Michigan State University, Ann Arbor, Michigan.

Lee E. Jacokes, PhD, is Professor of Psychology, Aquinas College, Grand Rapids, Michigan.

James F. Cassily, is Director, Neural Technology Research Center, 3090 Dawes, SE, Grand Rapids, Michigan 49508; cassily@interactivemetronome.com.

Stanley I. Greenspan, MD, is Clinical Professor of Psychiatry, Behavioral Sciences, and Pediatrics, George Washington University Medical School, Washington, DC.

Robert F. Tuchman, MD, is Professor of Neurology, University of Miami Medical School, Miami Children's Hospital, Miami, Florida.

Paul J. Stemmer, Jr., PhD, is Director, Center for Research, Madonna University, Livonia, Michigan.

This article was accepted for publication March 16, 2000.

The ability to attend, which begins early in life, is a vital part of the capacity to learn, concentrate, think, interact with others, and master basic academic skills (Greenspan, 1997; Greenspan & Lourie, 1981; Mundy & Crowson, 1997). Relative deficits in sustaining attention, inhibiting competing impulses, and engaging in joint attention can be found in attentional, learning, and developmental disorders. These deficits are part of several clinical disorders, including attention deficit disorder (ADD), pervasive developmental disorder (autistic spectrum disorders), language disorders, motor disorders, and specific learning disorders involving reading, math, and writing (Barkley, 1997a; Mundy, 1995).

Increasing evidence suggests that broad constructs such as motor planning and sequencing, rhythmicity, and timing are relevant to attentional problems. Barkley (1997b) postulated that deficits in inhibition and executive functions, which involve the regulation and sequencing of motor patterns and behavior, are important in understanding attention deficit hyperactivity disorder (ADHD). Additionally, several investigators have postulated important relationships between attention and aspects of motor regulation, including inhibition (Schonfeld, Shaffer, & Barmack, 1989), speed, rhythm, coordination, and overflow (Barkley, Koplowitz, Anderson, & McMurray, 1997; Denckla, Rudel, Chapman, & Krieger 1985; Piek, Pitcher, & Hay, 1999). Gillberg and Gillberg (1988) described a group of children with deficits in attention, motor control, and perception (termed DAMP syndrome), and in a recent study, Kadesjo and Gillberg (1998) found considerable overlap between attention deficits and motor clumsiness. In this group of children, the combination of both attentional and motor problems tends to worsen the prognosis (Hellgren, Gillberg, Bagenholm & Gillberg, 1994; Hellgren, Gillberg, Gillberg, & Enerskog , 1993). Piek et al. (1999) demonstrated that the severity of inattentive symptomatology in boys with ADHD is a significant predictor of motor coordination difficulties. Furthermore, recent work has suggested that approximately half of all children with developmental coordination disorder (DCD) have moderate to severe symptoms of ADHD, and a diagnosis of DCD at 7 years of age has been associated with restricted reading comprehension at 10 years of age (Kadesjo & Gillberg, 1999).

According to the Developmental, Individual-Difference, Relationship model (Greenspan 1992; Greenspan & Wieder, 1999), which uses dynamic systems theory (Gray, Kennedy, & Zemke, 1996a, 1996b; Smith & Thelen, 1993) to understand children's adaptive and maladaptive behavior, a child brings a variety of unique processing capacities, including motor planning and sequencing, into interactions with others and the physical environment in order to construct complex adaptive patterns, such as attending to and carrying out multistep actions in school and at home. Furthermore, considerable overlap exists in the neural networks involved in ADHD and the regulation of timing and the motor planning. These networks involve the prefrontal and striatal regions of the brain. A study using functional magnetic resonance imaging demonstrated that children with ADHD had subnormal activation of prefrontal systems responsible for high-order motor control (Rubia et al., 1999).

The relationship between motor regulation and attentional and executive functions suggests that technologies aimed at strengthening motor planning, sequencing, timing, and rhythmicity may have a role in improving the capacity to attend and learn (Greenspan, 1992). The Interactive Metronome®, a patented, PC-based interactive version of the traditional music metronome developed in 1992 (Cassily, 1996), is a new educational technology aimed at facilitating a number of underlying central nervous system processing capacities hypothesized to be involved in motor regulation. Noninteractive metronomes have been used as temporal teaching tools since being invented in 1696 by Étienne Loulié. The Interactive Metronome is the first to use the capabilities of modern computers to add an interactive element to this traditional tool. Instead of users having to rely on their own mental estimations of their temporal accuracy, the Interactive Metronome provides accurate (to .5 ms), real-time guide sounds to indicate users' temporal accuracy as they perform a series of prescribed movements. The tonally and spatially changing guide sounds enable users to deliberately correct their rhythmicity and timing errors as they are occurring.

Preliminary studies have shown that the level of a person's performance on the Interactive Metronome that involves planning, timing, and rhythmicity of motor regulation correlates with the severity of developmental, learning, and attentional problems, improvements in academic performance, and age-expected performance changes during the school years (Kuhlman & Schweinhart, 1999). Children with a range of developmental and learning problems in special education classes who trained on the Interactive Metronome have demonstrated gains in motor performance compared with a similar group without such training who demonstrated no gains over the same period (Stemmer, 1996).

Libkuman and Otani (1999) showed that Interactive Metronome training can improve motor control, focus, and athletic performance in golfers. The present study is the first controlled clinical trial of Interactive Metronome training on a group of children who meet the DSM-IV criteria for ADHD (American Psychiatric Association, 1994). The purpose of this study was to determine the effects of the Interactive Metronome on selected aspects of motor and cognitive skills in a group of children with ADHD.

Method

Design

This research used an experimental pretest–posttest measurement design (see Table 1). An Interactive Metronome treatment group was compared with a video treatment group and a traditional control group receiving no interventions.

Sample

Participants were drawn from the population of 6-year-old to 12-year-old boys with ADHD living in the greater metropolitan area in which the study was conducted. Seventy-

Table 1
Experimental Research Design

Treatment Group	Pretesting	Training	Posttesting
Interactive Metronome® (n = 19)	•	15, 1-hr sessions over a 3-week to 5-week period	•
Video game (n = 19)	•	15, 1-hr sessions over a 3-week to 5-week period	•
Control (n = 18)	•	None	•

five volunteers verified clinically as meeting DSM-IV criteria for ADHD by pediatricians, pediatric subspecialists, and psychologists or psychiatrists were recruited through local school districts, physicians, psychologists, psychiatrists, and advertisements in a local newspaper. All testing and treatments were given at no cost to the participants' parents. Test administrators screened, pretested, and posttested each child who was randomly assigned to them. All test administrators were paid, qualified psychometricians or licensed occupational therapists certified in administering their respective tests. Test administrators were not informed about the study's purpose and were blind to who received what treatment.

As a result of the screening, 19 boys were dropped from the volunteer pool because they either did not meet the clinical or research criteria or had severe learning, cognitive, neurological, anxiety, or depression problems. Demographically, the 56 qualified participants were 6 years to 12.5 years of age. Eighty-six percent were Caucasian, and 14% were of other races. Thirty-two percent had parents or guardians with annual incomes under $40,000, 38% from $40,000 to $69,000, and 30% with $70,000 or more. Eighty percent had parents or guardians with a college education.

Both parents and children were told that the purpose of the study was to "explore the use of nonpharmacological methods in the treatment and management of ADD and ADHD" and that the "treatments to be used in the study were interactive, computer-based treatment programs." No further information about the study was provided until completion of treatment and posttesting. One participant was belligerent toward his administrator and was removed from the study after the 2nd day. After completion of the study, the participants assigned to the video game and control groups received the Interactive Metronome treatment.

Instrument

Three major categories of performance were targeted for evaluation. The assessments were selected from those most commonly used by the psychological, occupational therapy, and educational communities. Only assessments that have been shown to be reliable and valid were used (see reference for each instrument). Summary and subtest scores from the following instruments were used to assess these categories of performance:

- *Attention and concentration:* (a) Tests of Variables of Attention (TOVA) is a 25-min computer-based assessment and one of the most widely used objective measures of ADHD (Greenberg & Dupuy, 1993). (b) Conners' Rating Scales–Revised (CRS-R), Teacher and Parent versions, is a questionnaire completed by the parents and teachers and one of the most widely used subjective measures of ADHD (Conners, 1990). (c) Wechsler Intelligence Test for Children–Third Edition is a well-known and widely accepted test of intelligence for children (Wechsler, 1992). (d) Achenbach Child Behavior Checklist is a questionnaire completed by parents that measures internalized problems and external behaviors (Achenbach & Edelbrock, 1991).
- *Clinical functioning:* (a) CRS-R. (b) Achenbach Child Behavior Checklist. (c) The Sensory Profile assesses auditory, visual, activity level, taste/smell, body/position, movement, touch, and emotional/social functioning (Dunn & Westman, 1995). (d) Bruininks-Oseretsky Test for Motor Efficiency (selected subtests) assesses bilateral coordination and upper-limb coordination, speed, and dexterity (Bruininks, 1978).
- *Academic and cognitive skills:* (a) Wide Range Achievement Test (WRAT 3) (reading and writing) assesses reading decoding, spelling, and math computation. (b) Language Processing Test assesses basic language (Wilkinson, 1993).

Participants were pretested and posttested at the same time of the day to control for medication schedules and circadian rhythms. On tests that offered equivalent forms, a different form was used for the posttesting than for the pretesting. The period between pretesting and posttesting was 4 to 5 weeks.

Test Administrators

The Interactive Metronome and video game group participants were randomly assigned to paid research administrators who treated participants of both groups. The administrators were college graduates, students, or persons without advanced degrees and with no previous formal therapy and teaching experience. Each administrator received 6 hr of instruction on both the Interactive Metronome and the video games.

Environments and treatment schedules for both groups were matched. Administrators followed a daily treatment regimen guide that controlled the structure of the sessions, time spent in conversation, and amount of encouragement given. Each participant was asked not to share his experiences with the other participants.

Procedure

Interactive Metronome group. The patented Interactive Metronome apparatus used in the study included a Pentium computer, the Interactive Metronome software program, two sets of headphones, and two contact-sensing triggers. One trigger, a special glove with a contact sensor attached to the palm side, sensed exactly when the triggered hand made contact with the other hand while clapping or when one hand was tapped on the thigh. The other trigger, a flat plastic pad placed on the floor, sensed when a toe or heel was tapped on it.

When the participant tapped a limb in time with the steady metronome reference beat sound heard in the headphones, the trigger sent a signal via a cable to the program. The Interactive Metronome analyzed exactly when in time the tap occurred in relation to the reference beat and instantaneously transposed the timing information into guidance sounds that the participant heard in the headphones as each tap occurred. The pitch and left-to-right headphone location of the guidance sounds precisely changed according to each tap's accuracy. The program-generated rhythmicity accuracy scores (Interactive Metronome scores), displayed in milliseconds on the screen, indicated to administrators how close in time the participant's responses were to the reference beat as they occurred. After each exercise, the participants were shown their scores. This feedback appeared to motivate them to do better.

The object of the Interactive Metronome treatment was to help participants improve their ability to selectively attend, without interruption by internal thoughts or external distractions, for extended periods. Simple limb motion exercises were used as systematic external catalysts to an underlying mental focus–improvement process. Each participant underwent 15, 1-hr Interactive Metronome treatment sessions, one session per day, spread out over a 3-week to 5-week period. Each session included 4 to 8 exercises that were repeated a specific number of times as prescribed in the daily treatment regimen guide. Exercises were done at a preset tempo of 54 repetitions/min, and the number of repetitions per exercise increased from 200 during the first session to a maximum of 2,000 during the ninth session.

The 13 Interactive Metronome treatment exercises were designed to help the participants put their efforts toward improving mental concentration rather than toward developing new physical motion techniques. The exercises included clapping both hands together, tapping one hand alone against the upper thigh, alternating toe taps on the floor, alternating heel taps on the floor, tapping one toe or heel alone on the floor, alternating between tapping one hand on the thigh and the toe on the floor, and balancing on one foot while tapping the other toe.

Before beginning their first Interactive Metronome treatment session, participants were given an automated Interactive Metronome pretest to quantify their ability to recognize timing patterns, selectively attend to a task, and make simple motion corrections. The pretest also indicated whether each participant had one or more rhythmicity deficiency patterns that needed to be addressed during their initial stage of treatment. Interactive Metronome treatment regimens were designed and accomplished in stages according to instructions in the daily treatment regimen guide.

During the first stage, the administrators helped the participants break the existing rhythmicity deficiency patterns that were identified during the pretest. The six rhythmicity deficiency patterns most frequently identified were the following

1. *Disassociative:* Three participants' responses were chaotic and random and not related to the beat in any way.
2. *Contraphasic:* Within a few beats, six participants' responses consistently moved to in between the beat rather than on the beat.
3. *Hyperballistic:* Sixteen participants used inappropriate snappy ballistic-type motions.
4. *Hyperanticipatory:* Eighteen participants' responses continually occurred much before the reference beat.
5. *Hypoanticipatory:* One participant's responses continually occurred much after the reference beat.
6. *Auditory hypersensitivity:* Seven participants were exceptionally distracted by the computer-generated guide sounds that were added to the headphone mix during the last test task, as indicated by their Interactive Metronome scores on that task, which were two to three times less accurate then those of the previous 13 tasks done without the guide sounds.

The initial Interactive Metronome treatment sessions were devoted to helping the participants learn how to discriminate between the sounds triggered by their own actions and the steady metronome beat. They were instructed to make smooth, controlled hand and foot motions that continuously cycled through a repeating pattern without stopping at any time between beats. Participants were repeatedly instructed to focus on the metronome beat and to try not to be interrupted by their own thoughts or things happening around them. When the participants had broken their existing rhythmicity patterns and were able to achieve the Interactive Metronome score average prescribed in the daily treatment regimen guide, they were considered to have achieved the adequate control and accuracy necessary to begin a second distinct phase of the Interactive Metronome treatment.

During the second treatment phase, participants were instructed to focus their attention only on the steady reference beat and ignore the computer-generated guide sounds, internal thoughts, and unrelated stimuli around them. They were also instructed to keep repeating their motion patterns without making any deliberate adjustments whatsoever. Doing so usually resulted in obvious improvements in the participant's Interactive Metronome score, and the entrainment experience of staying on beat without trying seemed to have a positive motivating effect. From session to session, participants increased the length of time they could selectively focus on the metronome beat without interruption, and their Interactive Metronome scores improved correspondingly. Most participants appeared to be highly motivated to achieve the best score possible during their Interactive Metronome training regimen. According to the Interactive Metronome scores, each participant improved his rhythmicity and was able to stay on task without being interrupted for significantly longer periods by the end of the training.

Video game group. Five commonly available PC-based, nonviolent video games were used as a treatment placebo for the video game group. Each game involved eye–hand coordination, advanced mental planning, and multiple task sequencing. In each game, the participant played against the computer, and at each new level achieved, the game became increasingly more difficult to play.

The test administrators followed a daily treatment regimen guide in the same manner as they did for the Interactive Metronome group. The prescribed video game exercises provided the participants with the same type of supervision, attention, and support as was received by the Interactive Metronome group. Each participant underwent 15, 1-hr video game training sessions, one session per day, spread out over a 3-week to 5-week period. Each training session involved a number of video game exercises, and the length of time they spent on each video game exercise typically increased from the first session to the last session.

Results

Sampling Design

After completion of pretesting of all 56 participants, a matched random assignment process was used to form the three treatment groups (i.e., Interactive Metronome, video game, control). Three factors were used in the matching process: medication dosage (mg/body weight), age, and severity of ADHD as measured by the TOVA. An analysis of variance (ANOVA) of these matching variables revealed no significant differences at the $p \leq .05$ level among the treatment groups. Chi-square analysis of three demographic variables—race, parental education, and parental household income—revealed no significant differences at the *p* £ .05 level, suggesting that the treatment groups were equal for these socioeconomic factors.

An ANOVA of the 58 pretest factors revealed only one significant difference among the treatment groups. Sakoda, Cohen, and Beall's (1974) table for tests of significant difference revealed the probability of this one significant difference in 58 significance tests occurring by chance to be $p > .50$, establishing this single occurrence to be likely a chance difference. The other 57 factors produced values in excess of $p > .05$, establishing the treatment groups' statistical equality.

Pattern Analysis

Pattern analysis of the 58 pretest factors examined the overall direction of mean differences between pretest and posttest phases for each group. In performing the analysis, the means for each test were computed, and the mean differences between the tests were determined. Each mean difference was dichotomized by whether the change represented an improvement or a decline in the desired direction for that test. For example, the posttest–pretest mean differences for the Wechsler Digit Span subtest for each treatment group were the following: Interactive Metronome = .473, control = –.278, and video game = –.054. The mean differences revealed improved performance in the Interactive Metronome group, whereas the control and video game groups showed decreased performance. Similar analyses were completed for all 58 test scores.

To statistically test the pattern, a binomial test was used to determine whether the proportion of dichotomous pairs (improvement vs. decline) was likely a chance occurrence (where the probability of either an improvement or decline = .50) or whether the directional proportion was so unusual as to reflect a non-chance event. The rationale for using a binomial test rests on the assumption that if a large number of variables collectively showed an unusual directional propensity (e.g., improved performance), this represented an overall pattern of change worthy of notice. The binomial test allows detection of a combined directional pattern that individual variables, taken one at a time, do not detect.

The pattern analysis revealed that the control group had 28 scores improve and 30 decline. Such a result has a high probable chance occurrence of $p = .8955$ and suggests that no significant combined directional pattern is present (Norusis, 1993). Analysis of the Interactive Metronome and video game groups produced significant improvement–decline patterns. For the Interactive Metronome group, 53 of the 58 variables showed improvement ($p \leq .0001$). For the video game group, 40 of 58 variables showed improvement ($p \leq .0058$). Both groups showed significant pattern increases in performance over the control group. The Interactive Metronome group experienced significantly better improvement than the video game group, suggesting that the Interactive Metronome treatment produced significant additional benefits above and beyond the experience of the video game and control group participants.

Significant Difference Analysis

The pattern analysis identified the overall improvement–decline characteristics of the test mean differences but did not address the magnitude of these differences. Because a pretest–posttest repeated measures design was used, an ANOVA for repeated measures (SPSS, 1988) was performed separately on each of the 58 variables. This approach was chosen to view the effects of the three treatment groups on each test score individually. However, one possible disadvantage of the approach is its potential of increasing Type 1 error.

Of the 58 test scores analyzed, 12 either had significant interaction effects (p = .0001–.047), suggesting that some combination of treatments and subgroup means were different, or there were significant pretest–posttest differences. Twelve significant differences out of 58 significance tests had a $p \leq .001$ at the .05 level of confidence (Sakoda et al., 1974), suggesting that these are not chance differences. Additionally, Keppel's (1973) calculation for the potential number of Type 1 errors over 58 separate experiments is 2.9. Thus, these 12 significant differences far exceed the calculated potential of 2.9 Type 1 errors, suggesting that these differences are real, significant differences.

Among the significant effects, seven significant differences between-phase effects were found (p =.0001–.023). This analysis finds the Interactive Metronome participants significantly improving their performance in identifying similarities and differences between concepts and in experiencing declines in aggressive behavior, as reported by their parents. Both the Interactive Metronome and video game treatments produced significant improvements on three Sensory Profile subtests, suggesting that both groups benefited from the attention and activities provided in these treatments. Parental reports on the Child Behavior Checklist also revealed significant declines in aggressive behavior for the Interactive Metronome group, a nonsignificant improvement for the video game group, and no improvement for the control group.

The remaining five tests had significantly different interaction effects (p = .0001–.047). These five tests were the WRAT 3 Reading subtest and four tests of the TOVA, including Omissions, RT (Response Time) Variability, Response Time Variability Total STD (Standard) Deviation, and ADHD Total Score. The significant interaction effects suggest that the posttest Interactive Metronome performances, though not significantly improved over the pretest performances, were significantly higher than the control and video game posttest performances. For all five tests, the patterns of differences were identical: Interactive Metronome performances improved, whereas both control and video game performances declined.

In summary, the pattern analysis revealed that both the Interactive Metronome and the video game groups experienced significant improvement patterns across the 58 test scores. Additionally, the Interactive Metronome group had a significantly stronger improvement pattern than the video game group, showing improvements over 53 test scores compared with 40 for the video game group. This finding supports the hypothesis that Interactive Metronome training produced a stronger improvement pattern than the video game group for boys with ADHD.

Analysis of test means found 12 factors with significant quantitative changes among the various group and treatment combinations. The Interactive Metronome group showed significant pretest–posttest improvement in identifying similarities and differences and reduction of aggression problems compared with the other two treatment groups. Both the Interactive Metronome and the video game groups showed significant improvements in three sensory processing tasks and in parental reports of impulsiveness and hyperactivity. Only parents of the Interactive Metronome participants, however, rated their children as significantly less aggressive ($p \leq .001$) after the treatment period. Additionally, five tests measuring reading and four characteristics of attention revealed that the Interactive Metronome group had significantly higher posttest performances than the other two groups.

Discussion

The results indicated that boys with ADHD who received the Interactive Metronome intervention improved significantly more in areas of attention, motor control, language processing, reading, and ability to regulate aggression than boys receiving either the video game treatment or no treatment. Participants who received video game treatment improved more than the participants in the control group on a number of measures as well, demonstrating that focused perceptual activities and support alone may be helpful for selected areas of functioning. The video game group, however, showed decreased performance in selected areas involving modulation and control, such as consistency of concentration, reaction time, and overall attention.

Interactive Metronome treatment, on the other hand, only showed improved performance, including significant positive gains, over the video game treatment on a series of TOVA attentional tasks measuring lack of errors and distractibility, consistency of reaction time, and overall attention; selected language (i.e., similarities and differences); academic tasks (reading); and control of aggression. In addition, pattern analysis was used to control for the effect of using a large number of assessments and demonstrated that the differences between the group patterns were significant. The National Institutes of Health (NIH, 1997) asserted that studies on ADHD interventions must properly control for the positive overall effect of attentive adult interaction, alone. Consistent with NIH guidelines, two of the three groups in this study received adult attention during the treatment period.

Limitations

Only male participants in a defined age range were included to minimize age and gender variation, thereby limiting generalizability to the other gender and age groups. The variables measured by the assessments are limited to selected aspects of attention, motor control, language, cognition, and learning.

In this study, Interactive Metronome training influenced a number of performance capacities. A possible explanation for the positive changes is the central role of motor planning and sequencing in each performance area.

In a dynamic systems model (Smith & Thelen, 1993), critical variables, such as the ability to plan and sequence actions, may influence a broad array of adaptive functions, including attention (Greenspan, 1992).

Directions for Future Research

The results of the current study suggest directions for further research, including replications of the current study on larger populations (which might permit the identification of characteristics associated with different patterns of response to metronome training), on girls, and on more socioeconomically diverse populations to observe potential components of different environmental contexts. Further research could also investigate subgroups that are based on both metronome performance and the child's processing profile.

Specific variations of the Interactive Metronome training process also need exploring, including increasing the number of sessions and overall repetitions, timing accuracy goals, and varying length of follow-up time to observe stability of the treatment effect. In addition, further research is needed to understand more fully both the dynamic systems and the underlying central nervous system mechanisms involved in motor regulation as well as the way in which Interactive Metronome training influences these processes. The Interactive Metronome may be the first technology that can allow the creation of a database and classification of "timing" to help compare the effects of interventions that influence timing in a variety of perceptual-motor processes.

Conclusion

From a dynamic systems perspective (Gray et al., 1996a, 1996b; Smith & Thelen, 1993), many processes, including the timing and rhythmicity of motor behavior, influence motor planning. In turn, motor planning interacts with other factors, including learning opportunities and environmental demands, to influence patterns of self-regulation and functioning at home, in school, and with peers. Until recently, interventions to strengthen these capacities have been limited to working with overt or surface behavior in educational or therapeutic settings. The present study suggests that Interactive Metronome training can improve aspects of attention, motor, and perceptual-motor functioning; cognitive and academic performance; and the control of aggression in children with major attentional problems. Hence, Interactive Metronome training may complement existing interventions for these children. ▲

References

Achenbach, T. M., & Edelbrock, C. E. (1991). *Child Behavior Checklist.* Burlington, VT: University of Vermont.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Barkley, R. (1997a). Attention-deficit/hyperactivity disorder, self-regulation, and time: Toward a more comprehensive theory. *Journal of Developmental and Behavioral Pediatrics, 18,* 271–279.

Barkley, R. A. (1997b). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121,* 65–94.

Barkley, R. A., Koplowitz, S., Anderson, T., & McMurray, M. B. (1997). Sense of time in children with ADHD: Effects of duration, distraction, and stimulant medication. *Journal of the International Neuropsychological Society, 3,* 359–369.

Bruininks, R. H. (1978). *Bruininks-Oseretsky test of motor proficiency: Examiners manual.* Circle Pines, MN: American Guidance Service.

Cassily, J. F. (1996, June 25). Methods and apparatus for measuring and enhancing neural motor coordination [U.S. patent 5,529,498].

Conners, C. K. (1990). *Conners' Rating Scales–Revised.* Lutz, FL: Psychological Assessment Resources.

Denckla, M. B., Rudel, R. G., Chapman, C., & Krieger, J. (1985). Motor proficiency in dyslexic children with and without attentional disorders. *Archives of Neurology, 42,* 228–231.

Dunn, W., & Westman, K. (1995). *Sensory Profile study: Occupational therapy education.* Kansas City, KS: University of Kansas.

Gillberg, I. C., & Gillberg, C. (1988). Children with deficits in attention, motor control and perception (DAMP): Need for specialist treatment. *Acta Paediatrica Scandinavica, 77,* 450–451.

Gray, J. M., Kennedy, B., & Zemke, R. (1996a). Application of dynamic systems theory to occupation. In R. Zemke & F. Clark (Eds.), *Occupational science: The evolving discipline* (pp. 300–324). Philadelphia: F. A. Davis.

Gray, J. M., Kennedy, B., & Zemke, R. (1996b). Dynamic systems theory: An overview. In R. Zemke & F. Clark (Eds.), *Occupational science: The evolving discipline* (pp. 297–308). Philadelphia: F. A. Davis.

Greenberg, L. M., & Dupuy, T. R. (1993). *Interpretation manual for the Test of Variables of Attention computer program.* (Available: Universal Attention Disorders, 4281 Katella Avenue, #215, Los Alamitos, CA 90720)

Greenspan, S. I. (1992). *Infancy and early childhood: The practice of clinical assessment and intervention with emotional and developmental challenges.* Madison, CT: International University Press.

Greenspan, S. I. (1997). *The growth of the mind and the endangered origins of intelligence.* Reading, MA: Addison-Wesley Longman.

Greenspan, S. I., & Lourie, R. S. (1981). Developmental structuralist approach to the classification of adaptive and pathologic personality organizations: Application to infancy and early childhood. *American Journal of Psychiatry, 138,* 725–735.

Greenspan, S. I., & Wieder, S. (1999). A functional developmental approach to autism spectrum disorders. *Journal of the Association for Persons With Severe Handicaps, 24,* 147–161.

Hellgren, L., Gillberg, I. C., Bagenholm, A., & Gillberg, C. (1994). Children with deficits in attention, motor control and perception (DAMP) almost grown up: Psychiatric and personality disorders at age 16 years. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 35,* 1255–1271.

Hellgren, L., Gillberg, C., Gillberg, I. C., & Enerskog, I. (1993). Children with deficits in attention, motor control and perception (DAMP) almost grown up: General health at 16 years. *Developmental Medicine and Child Neurology, 35,* 881–892.

Kadesjo, B., & Gillberg, I. C. (1998). Attention deficits and clumsiness in Swedish 7-year-old children. *Developmental Medicine and Child Neurology, 40,* 796–804.

Kadesjo, B., & Gillberg, I. C. (1999). Developmental coordination disorder in Swedish 7-year-old children. *Journal of the American Academy of Child and Adolescent Psychiatry, 38,* 820–828.

Keppel, G. (1973). *Design and analysis: A researcher's handbook.* Englewood Cliffs, NJ: Prentice Hall.

Kuhlman, K., & Schweinhart, L. J. (1999). *Timing in child development.* Ypsilanti, MI: High/Scope Educational Research Foundation.

Libkuman, T., & Otani, H. (1999). *Training in timing improves* *accuracy in golf.* Unpublished manuscript, Central Michigan University, Mt. Pleasant, MI.

Mundy, P. (1995). Joint attention and social-emotional approach behavior in children with autism. *Development and Psychopathology, 7,* 63–82.

Mundy, P., & Crowson, M. (1997). Joint attention and early social communication: Implications for research on intervention with autism. *Journal of Autism and Developmental Disorders, 27,* 653–676.

NIH consensus statement. (1997). *Archives of General Psychiatry, 54,* 865–870.

Norusis, M. (1993). *SPSS for Windows base system user manual* (Release 6.0). Chicago: Statistical Package for the Social Sciences.

Piek, J. P., Pitcher, T., & Hay, D. A. (1999). Motor coordination and kinaesthesis in boys with attention deficit-hyperactivity disorder. *Developmental Medicine and Child Neurology, 41,* 159–165.

Rubia, K., Overmeyer, S., Taylor, E., Brammer, M., Williams, S., Simmons, A., & Bullmore, E. (1999). Hypofrontality in attention deficit hyperactivity disorder during higher-order motor control: A study with functional MRI. *American Journal of Psychiatry, 156,* 891–896.

Sakoda, J. M., Cohen, B. H., & Beall, G. (1974). Test of significance for a series of statistical tests. *Psychological Bulletin, 51,* 172–175.

Schonfeld, I., Shaffer, D., & Barmack, J. (1989). Neurological soft signs and school achievement: The mediating effects of sustained attention. *Journal of Abnormal Child Psychology, 17,* 575–596.

Smith, L. B., & Thelen, E. (Eds.). (1993). *A dynamic systems approach to development: Applications.* Cambridge, MA: MIT Press.

Statistical Package for the Social Sciences. (1988). SPSS-X users guide (3rd ed.). Chicago: Author.

Stemmer, P. M. (1996). *Improving student motor integration by use of an Interactive Metronome.* Paper presented at the 1996 Annual Meeting of the American Educational Association, Chicago.

Wechsler, D. (1992). *Wechsler Intelligence Scale for Children manual* (3rd ed.). San Antonio, TX: Psychological Corporation.

Wilkinson, G. S. (1993). *Wide Range Achievement Test administration manual.* Wilmington, DE: Wide Range.

Psychology in the Schools, Vol. 44(8), 2007
Published online in Wiley InterScience (www.interscience.wiley.com).

DOI: 10.1002/pits.20270

IMPROVEMENTS IN INTERVAL TIME TRACKING AND EFFECTS ON READING ACHIEVEMENT

GORDON E. TAUB

University of Central Florida

KEVIN S. MCGREW

Institute for Applied Psychometrics

TIMOTHY Z. KEITH

University of Texas

This study examined the effect of improvements in timing/rhythmicity on students' reading achievement. 86 participants completed pre- and post-test measures of reading achievement (i.e., Woodcock-Johnson III, Comprehensive Test of Phonological Processing, Test of Word Reading Efficiency, and Test of Silent Word Reading Fluency). Students in the experimental group completed a 4-week intervention designed to improve their timing/rhythmicity by reducing the latency in their response to a synchronized metronome beat, referred to as a synchronized metronome tapping (SMT) intervention. The results from this *non-academic* intervention indicate the experimental group's post-test scores on select measures of reading were significantly higher than the non-treatment control group's scores at the end of 4 weeks. This paper provides a brief overview of domain-general cognitive abilities believed effected by SMT interventions and provides a preliminary hypothesis to explain how this *non-academic* intervention can demonstrate a statistically significant effect on students' reading achievement scores. © 2007 Wiley Periodicals, Inc.

In recent years the role of the school psychologist has expanded to include greater involvement in students' reading acquisition, performance, and curriculum-based evaluation. This increased participation may be attributed to several national initiatives including Reading First under No Child Left Behind (U.S. Department of Education, 2002), the National Reading Panel's (2000) report, the Individuals with Disabilities Education Improvement Act (2004), and the impact of empirical research in reading on district- and state-level policies and procedures (e.g., Daly & McCurdy, 2002; Sheridan, 2004). Recent technological advancements also provided school psychologists with a broader understanding of the process of reading at a physiological level. Results from neuroscience studies (e.g., functional magnetic resonance imaging investigations involving individuals experiencing reading difficulties or diagnosed with dyslexia) have provided new insights into the process of reading at the neural level (e.g., see Katzir & Paré-Blagoev, 2006). This groundbreaking research has demonstrated individual differences in the functions of anatomically similar brain regions of impaired readers and nonimpaired readers (Katzir & Paré-Blagoev, 2006; Shaywitz & Shaywitz, 2005; Shaywitz et al., 1999, 2003).

The integration of our understanding of the process of reading at a physiological level with reading at a behavioral level (i.e., neuroscience-based interventions) may be the next frontier for school psychologists and reading research. One intervention that has received considerable empirical attention, both pro and con, is the FastForWard method (Tallal, Miller, Jenkins, & Merzenich, 1997). A lesser known neuroscience-based intervention is the use of synchronized metronome tapping, which links research on mental interval timekeeping (e.g., see Buhusi & Meck, 2005) and academic achievement. Preliminary results from this research indicate that children diagnosed

This study was funded in part through a grant from the Interactive Metronome, Inc.
Correspondence to: Gordon E. Taub, University of Central Florida, ED 123K, Orlando, FL 32816. E-mail: gtaub@mail.ucf.edu

with dyslexia may have deficiencies in their timing and rhythm abilities, as evidenced by their responding within a wider range of times on either side of a metronome beat, when compared to nonimpaired readers (Wolff, 2002). Similarly, McGee, Brodeur, Symons, Andrade, and Fahie (2004) reported children diagnosed with a reading disability differed from children diagnosed with attention-deficit/hyperactivity disorder (ADHD) on retrospective time perception, a finding interpreted as consistent with Barkley's (1997) behavioral inhibition theories. Research also implicated mental or interval timekeeping (time perception) in a number of academic and behavioral disorders (see McGee et al., 2004). Some researchers believe the connection between timing/rhythm and reading may be so robust that a student's mean latency response to a metronome beat may predict performance on standardized reading tests (Waber et al., 2003; Wolff, 2002). Furthermore, a recent study has suggested that elementary timing tasks may represent a form of *temporal g* that is more strongly correlated ($r = .56$) with psychometric *g* than the standard *reaction time g* ($r = -.34$) approach to measuring the *essence* of general intelligence (Rammsayer & Brandler, in press). Given the growing evidence suggesting a potentially important link between mental interval timekeeping and cognition and learning (Buhusi & Meck, 2005; Rammsayer & Brandler, in press), the connection between timing-based neuroscience interventions (e.g., synchronoized metronome tapping) and academic achievement warrants investigation.

To investigate the relationship between improvements in timing and rhythm (due to synchronized metronome tapping-based intervention) on reading achievement, Taub, McGrew, & Lazarus (2007) administered subtests from the Woodcock-Johnson Tests of Achievement III (WJ-III ACH; Woodcock, Mather, & McGrew, 2001) as pre- and posttest measures of reading. In this study, over 250 high-school-aged participants were randomly assigned to either a control or experimental group. The experimental group participated in a rhythmic synchronization metronome-based assessment and intervention technique (herein after referred to as the Interactive Metronome [IM] method), a *nonacademic* intervention. The IM treatment sessions lasted for approximately 45 minutes each day for total of about 15 hours. (The IM intervention method will be discussed in detail below.) The results from this study indicated, when compared to the control group, the experimental group demonstrated statistically significant improvements on the WJ-III ACH posttest measures of broad reading and reading fluency. Participants who received IM-based interventions also demonstrated statistically significant improvements in domains other than reading.

IM training was also reported to produce positive effects in a number of nonacademic domains. For example, after receiving IM training, participants demonstrated statistically significant improvements in golf performance (Libkuman & Otani, 2002). Shaffer et al. (2001) reported that boys prediagnosed with ADHD demonstrated improved performance, when compared to two ADHD control groups, in the domains of attention, language processing, motor control, reading, and parent report of regulation of aggressive behavior after their participation in an IM-based intervention.

Mental Interval Timing Research and Models

Cognitive psychology's interest in mental timekeeping has spanned decades. For example, cognitive differential psychologists first reported the identification of a *temporal tracking* capability in 1980 (Stankov, Horn, & Roy, 1980). Temporal tracking was identified as being found in various auditorily presented tasks that involved the mental counting or rearrangement of temporal sequential events (e.g., reorder a set of musical tones; Carroll, 1993).

Researchers in cognitive psychology have studied the phenomenon of *interval timing* through a number of research paradigms, one which requires individuals to maintain synchrony (via a bimanual motor response) with auditory tones (e.g., from a metronome), also known as synchronized metronome tapping (SMT). Tapping in synchrony with a metronome requires an individual to correct for asynchronies in their response to a reoccurring beat. The most viable theoretical explanation for SMT behavior can be derived from the pacemaker-accumulator model, which is based on scalar timing/expectancy theory (see Buhusi & Meck, 2005). Briefly, SMT asynchrony corrections are thought to be accomplished through an internal adjustment to the phase of one's underlying master mental time clock (Buhusi & Meck, 2005; Vorberg & Fuchs, 2004). This error correction is triggered when observed temporal deviations (as determined via the accumulation, in a short-term storage *accumulator*, of neural pulses or tics from a cognitive *pacemaker*) are determined to differ from a reference *standard* (which is maintained in a *reference memory*), via performance feedback. This process is referred to as an *automatic phase adjustment*. The allocation of *attentional resources* and the minimization of stimuli that may divert cognitive processing resources away from timing have been hypothesized to play a significant role in mental interval timekeeping and metronome-based synchronization of rhythmic movements (Brown & Bennett, 2006; Buhusi & Meck, 2005). In addition, the quickness and efficiency of the phase adjustment mechanism is believed to eliminate the necessity for, or excessive reliance on, long-term memory (e.g., accessing the reference memory) or learning (Vorberg & Fuchs, 2004).

How SMT-Based IM Training Works

During IM training participants wear a headphone and listen to a reoccurring metronome beat. As they listen to the beat, they engage in physical movements such as clapping hand-to-hand with a sensor on one palm as they match their physical movement to the presentation of the beat (e.g., clap at the beat). The goal of IM training is to reduce the mean negative synchronization error during normal tracking of the regularly occurring metronome beat (clapping prior to or past the beat).

During training, participants receive feedback through an auditory guidance system as they progress through the simple, interactive physical movements. Although feedback is also provided through visual stimuli, the auditory feedback guidance system is the primary feedback method. The auditory feedback system provides tonal stimuli that indicate whether the participant responded *prior to*, *at*, *or past* the regularly occurring auditory metronome beat. The accuracy of participants' expectancy response to the metronome beat is provided in milliseconds (ms), with different tones indicating *far from*, *close to*, or *at* the metronome beat. A visual reading of millisecond latency is also presented on a computer screen.[1] The purpose of IM training is to improve participants' timing/rhythmicity by reducing the latency between the onset of the metronome beat and participant's expectancy response to the beat. After about 3–4 weeks of training, or 15–18 hours, participants are typically able to respond to within approximately 15 ms on either side of the beat. This compares to the average 80–100 ms latency response prior to training. At the completion of training, participants typically have engaged in approximately 25,000 motoric repetitions. These movements are the physical indication of one's expectancy of the onset of the metronome beat. Collectively, results from initial studies suggest that statistically significant improvements in a *domain-specific* SMT-based intervention are associated with statistically significant *domain general* improvements in the areas of academics, ADHD, and sports. How can rhythmic SMT-based interventions result in improved performance across such diverse domains of human performance as academics, ADHD, golf, and tennis?

[1]Readers are referred to the Interactive Metronome, Inc.'s Web site to view a corporate-sponsored video showing IM training or to obtain additional information: http://www.interactivemetronome.com.

Purpose

Although hypothesized domain-specific cognitive mechanisms are possible, the domain-general or cross-domain SMT training effect is intriguing and argues first for replication of prior studies and second for investigation of potential domain *general* cognitive mechanisms to account for observed cross-domain improvements. Given this assumption, the purpose of this study was twofold.

The first purpose was to replicate an earlier study by examining the impact of improvements in timing/rhythmicity on students' reading achievement. The second purpose was to offer preliminary hypotheses that will contribute to a better understanding of the across-domain general cognitive mechanisms that may explain SMT treatment effects across such diverse human performance domains as academics, ADHD, and sports.

Method

Participants

Study participants included 86 students attending a public charter school receiving Title 1 funding located in Central Florida. As a public charter school, the school is a part of the public school system; the key difference between the public charter school and a public school is that the charter school receives funding directly from the State of Florida. The school currently has 133 students and provides education from kindergarten through fifth grade. All students attending the school are African-American, and 83% of the students receive free lunch. The study participants ranged in grade from first to fourth grade. There were 16 first-, 36 second-, 23 third-, and 11 fourth-grade students in the study. A total of 37 participants were male and 48 were female. Participants' ages ranged from 7 years old to 10 years old with a mean of 8.15 years ($SD = 1.0$).

Instruments

The instruments administered to evaluate the effects of IM training on participants' academic achievement and attention/concentration include selected subtests of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999), Test of Silent Word Reading Fluency (TOSWRF; Mather, Hammill, Allen, & Roberts, 2004), Test of Word Reading Efficiency (TOWRE; Torgesen, Wagner, & Rashotte, 1999), and the WJ-III ACH (Woodcock et al., 2001). Table 1 provides a brief description of each test and identifies the specific subtests administered from each instrument.

Reliability

Most of the reported average internal consistency and alternate form reliability coefficients of the CTOPP exceed .80 and the test–retest coefficients range from .70 to .92 (Wagner, Torgesen, & Rashotte, 1999). The reported average alternate forms' reliability coefficients of the TOWRE all exceed .90 and the test–retest coefficients range from .83 to .96 (Torgesen, Wagner, & Rashotte, 1999). The median reliability coefficients of the tests selected from the WJ-III ACH are all at or above .87 (McGrew & Woodcock, 2001).

A lesser know test was the Test of Silent Word Reading Fluency. This instrument was standardized on 3592 individuals representing demographic characteristics that were similar to the 2001 U.S. Census data in terms of geographic region, gender, race, ethnicity, and parents' educational background. The instrument's normative tables are grouped in 3-month intervals for students ages 6-6 through 7-11, 6-month intervals for students 8-0 through 10-11, and at 1-year intervals for students ranging from 11-0 through 17-11 years of age. Reported test–retest reliabilities

Table 1
Names and Description of the Pretest and Posttests

Test	Description of tests and combinations of tests
Test of Oral Word Reading Efficiency	*Sight Word Efficiency*: A timed test of word recognition and decoding fluency, measures the ability to accurately and quickly recognize familiar words *Phonemic Decoding Efficiency*: A timed test measuring the ability to accurately and quickly read phonetically regular nonsense words. *Total Word Reading Efficiency*: Combines Sight Word Efficiency and Phonemic Decoding Efficiency.
Test of Silent Word Reading Fluency	Students are presented with several rows of words, which increase in difficulty. There are no spaces between the words (e.g., didhimgot). Students are required to draw a line between the boundaries of as many words as possible (e.g., did/him/got) within a 3-min time limit.
The Comprehensive Test of Phonological Processing	*Blending Nonwords*: Phonetic coding synthesis task of nonwords—an auditory processing task that is independent of acquired knowledge (less dependent on students' existing knowledge). *Segmenting Nonwords*: Phonetic coding analysis task of nonwords—an auditory processing task that is independent of acquired knowledge. *Rapid Digit Naming*: Rapid automatized naming test of digits. *Rapid Letter Naming*: Rapid automatized naming test of letters. *Rapid Naming Composite*: Combines Rapid Digit Naming and Rapid Letter Naming. *Alternate Phonological Awareness Composite*: Combines Blending Nonwords and Segmenting Nonwords.
Woodcock-Johnson III Tests of Achievement	*Letter-Word Identification*: Untimed measure of sight-word recognition. *Passage Comprehension*: Measure of reading comprehension and word knowledge. *Reading Fluency*: A timed test measuring reading speed, automaticity and rate of test taking. *Word Attack*: Untimed test requiring pronouncing nonwords that conform to English spelling rules.

for students ranging in age from 7 to 10 years of age, the age range of the present study, were all above .80, and the alternate form reliability coefficients exceeded .85 (Mather et al., 2004).

Procedure

All students completed a pretest battery of psychoeducational instruments (see Table 1). After completing the pretests, students were randomly assigned to either an experimental or control group. The experimental group participated in the IM intervention, at their school, during regular school hours. While the experimental group was participating in the IM intervention, the control group and nonparticipating classmates engaged in recess activities. Students in the experimental group were divided into four groups, one for each grade level. Two certified master trainers worked separately with each of the four grade-level groups. The groups ranged in size from 7 to 12 participants. The students in the experimental group participated in an average of 18 sessions, each lasting approximately 50 minutes. There was one treatment session each day per group. Upon completion of the IM intervention, posttests were administered to all participants. The same tests were used during the pre- and posttest administrations.

Participants completed both individually and group administered tests; however, the TSWRF and WJ-III ACH's Reading Fluency were the only group-administered tests. During the individual assessment each evaluator worked with a student one on one. The individual assessment took approximately 35 minutes to complete. Group administrations were conducted in the students' own classrooms and participants from the experimental and control group completed all group tests together as classmates. Students who were unable to participate and/or who were absent on the day of the group assessments completed the group tests either individually or with other nonclassmate students. During all test administrations the test proctors and administrators were unaware of each student's group assignment. A lead test administrator directed all group assessments. The administrator followed the standardized instructions included in each test's manual. For one test, WJ-III ACH Reading Fluency, minor modifications were made in standardized administration procedures to facilitate group administration of the test. Several steps were followed to ensure that standardized test administration procedures were followed as closely as possible. These steps included (a) a doctoral-level proctor was present during all group administrations, (b) a minimum of one proctor to every four students was maintained during all group administrations, (c) all test proctors were graduate-level school psychology students who either completed or were near completion of their second psychoeducational assessment course, and (d) if a student did not accurately complete a sample item, the group administration was stopped and the proctor followed standardized administration procedures to ensure adequate completion of the sample item. All students progressed through the group test administration at the same time.

Results

Unless otherwise noted, all analyses controlled for pretest scores using the same measure as the posttest (through analysis of covariance). For analyses that did not use developmentally based scores, such as raw or growth scores, age was also controlled in the analyses by entering age as a covariate in the ANCOVA. Given the prediction that statistically significant differences would favor the experimental group, one-tailed tests ($\alpha = .05$) were used to evaluate statistical significance.

Effects on Timing/Rhythm

The initial analysis examined the effect of IM training on timing and rhythm as measured by the IM assessment system. The IM treatment had a statistically significant effect on posttest timing and rhythm scores, with pretest score controlled, $F(1,76) = 107.376$, $p < .001$. Furthermore, the treatment had a large effect (Thompson, 1999) on the posttest outcome ($\eta^2 = .586$, $g = 1.974$). IM training accounted for more than 50% of the variance in IM posttest scores and resulted in close to a two standard deviation increase in those scores (with IM pretest scores controlled).

It seems likely that IM training should be more effective for children who initially showed poor performance (high scores) on the measure of timing and rhythm. Sequential multiple regression was used to evaluate the possibility of a statistically significant interaction between the pretest and treatment. The IM posttest was regressed on the centered IM pretest and group membership in one block, with the centered pretest by group cross-product entered in a second block. As summarized in Table 2, the addition of the cross-product to the regression resulted in a statistically significant increase in R^2, indicating that the Pretest $\times$ Treatment Group interaction was statistically significant. The nature of the interaction is demonstrated in Figure 1, which shows separate regression lines for the posttest on the pretest, by treatment group. The lines show that the experimental group performed better on the posttest than did the control group, but that training was indeed most effective for participants with poor initial timing/rhythmicity.

Reading

Multivariate analysis of covariance (MANCOVA) was used to test the effect of IM training on the four measures of reading skill from the WJ-III ACH (Letter-Word Identification (LW-ID),

Table 2
Sequential Multiple Regression to Test Whether IM Training Was More Effective for Those with Initially High (Poor) Scores on Timing/Rhythmicity

Variables entered	ΔR^2	p
IM Pretest (centered), Treatment Group	.707	$<.001$
Pretest by group cross-product	.082	$<.001$

Reading Fluency, Passage Comprehension, and Word Attack). Pretest scores on these measures were used as covariates. As recommended by the test authors, *W* scores (a continuous, equal interval growth scale scores) were used for these analyses. The results of this analysis (and subsequent MANCOVA results) are summarized in Table 3. As shown in the table, the IM training did not demonstrate a statistically significant effect on reading achievement as measured by the WJ-III achievement tests.

Table 3 also shows the effects of IM training on measures of reading efficiency, TOWRE (Sight Word, Phonemic Decoding), and fluency, TSWRF. For this set of analyses, standard scores (M = 100, SD = 15) were used as both pre- and posttest scores; pretest scores and age were the

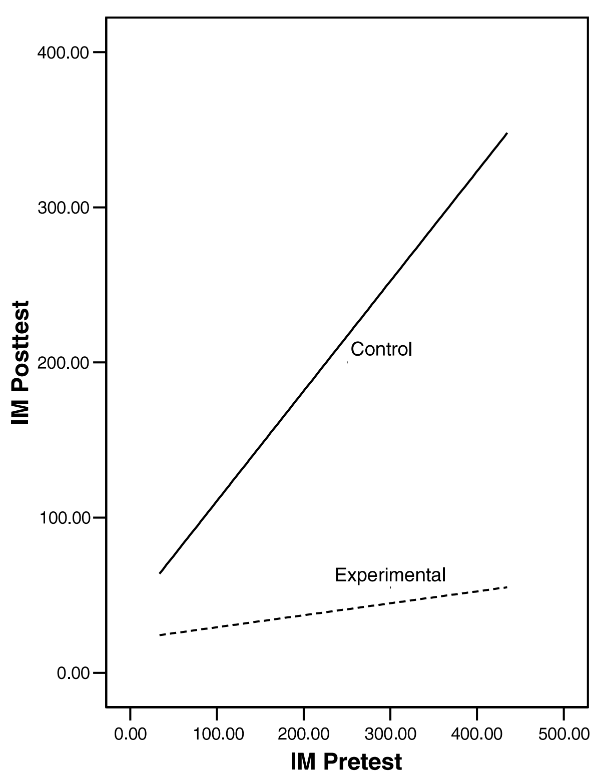


FIGURE 1. Interaction between IM pretest and IM training. The regression lines show that IM training was most effective in improving the timing and rhymicity of children with initial poor performance (low scores represent better performance).

Table 3
MANCOVA Results: Effect of IM Training on Reading

Measures	Hotelling's trace	$F(df)$	p	η^2
WJ Achievement Reading	.045	.842 (4, 75)	>.05	.043
Reading Efficiency & Fluency	.098	2.414 (3, 75)	.037	.089
CTOPP Phonological Processing	.205	3.899 (4, 76)	.003	.170

controlled variables. As shown in the table, the IM training produced a statistically significant effect on measures of reading efficiency and fluency. Participants who received IM training scored at a higher level on the multivariate dependent variable. The IM treatment accounted for 8.9% of the variance in reading efficiency and fluency, a small effect size (Keith, 2006, p. 508). Follow-up tests (univariate ANCOVAs) revealed a statistically significant effect for the TOWRE Sight Word Efficiency measure, $F(1,76) = 5.881$, $p = .009$, $\eta^2 = .072$, $g = .481$,[2] but not for the other measures.

Table 3 also shows the results of analyses of the IM effects on phonological processing skills as measured by the CTOPP (digit naming, letter naming, segmenting, and blending). Participants who received IM training demonstrated statistically significantly higher CTOPP scores, and the IM treatment accounted for 17% of the variance in CTOPP scores, a moderate effect. Univariate follow-up statistical analyses revealed statistically significant effects on the letter naming subtest, $F(1,79) = 8.680$, $p = .002$, $\eta^2 = .099$, $g = .536$, but not for the other components of the CTOPP.

Discussion

The current study employed a pre-/posttest evaluation design to investigate the effect of a specific SMT intervention (viz., Interactive Metronome) on reading performance in a sample of 86 first-, second-, third-, and fourth-grade students in a public charter school receiving Title 1 funding. Participants were randomly assigned to either an experimental (IM) or control group. The experimental group participated in a 3–4-week IM intervention designed to improve their timing/ rhythmicity. The control group engaged in recess activities with nonparticipating classmates during each of the approximately 50 minute daily intervention sessions. All participants completed the same reading pre- and posttest measures, which were then analyzed via statistical methods that controlled for initial pretest performance levels and age (ANOCOVA, MANOVA).

Timing and Rhythmicity Treatment Findings

The results indicated that the IM treatment produced significant improvements in the timing and rhythmicity of elementary school students (as measured by the IM measurement system). The students in the IM treatment group, when compared to the control group, demonstrated statistically significant improvements, close to a two standard deviation increase in measured timing and rhythmicity scores.

IM treatment transfer effects were evaluated vis-à-vis pre-/posttest changes on standardized measures of reading achievement. The reading-dependent variables sampled four of the five reading skills identified as critical for early reading success by the National Reading Panel (2000). The

[2] We know of no formula for calculating Hedges' g for overall MANOVA results. Therefore, partial η^2 is reported for MANOVA results and both η^2 and g are reported for the univariate follow-up tests.

reading-dependent variables included standardized measures of phonics, phonological awareness, reading fluency, and comprehension. The fifth key reading skill, vocabulary, was not measured.

Before discussing the IM academic transfer effect findings, it is important to note this intervention did *not* include instruction or training of any kind in phonics, phonological awareness, and/or reading—this was *not* an *academic* intervention. The IM intervention is designed to improve participants' timing and rhythmicity through beeps, tones, tapping, and clapping. In other words, it would not be expected that participants in an intervention designed to improve timing and rhythmicity would demonstrate changes in reading achievement. Furthermore, the experimental IM treatment lasted approximately 3–4 weeks. Developmental *growth* curves based on nationally standardized reading tests (McGrew & Woodcock, 2001) suggest that similarly aged students (8.2 years) typically demonstrate little academic growth (as reflected by norm-referenced tests) over a 3–4-week period.

Reading Achievement Findings

Analysis of the individual reading tests indicated that the IM intervention produced significant transfer effects in phonics, phonological awareness, and reading fluency. Students in the IM experimental group demonstrated statistically significant improvement in their ability to *fluently* recognize familiar words within a *limited timeframe* (TOWRE test). In contrast, no significant treatment effect was demonstrated on an *untimed* word recognition measure (WJ-III LW-ID test). It is important to note that the primary difference between the TOWRE and WJ-III LW-ID tests is that of a *rate fluency* (TOWRE) versus *level* (WJ-III LW-ID) distinction. *Rate fluency* refers to the time taken to work from the beginning of a test to the end of a test. *Level* refers to the difficulty of an item or task (see Carroll, 1993).

Within the context of a rate-fluency/level-ability distinction, the current results suggest the hypothesis that although students did not *learn* to recognize more familiar words in isolation (i.e., their absolute word recognition *level* did not increase), they were able to recognize the words they previously *knew* faster (i.e., the fluency of their level of word recognition skills was improved). It appears that SMT-based IM treatments may demonstrate transfer effects on reading fluency/efficiency of existing word recognition skills, but not increase the overall level of word recognition skills in a student's repertoire.

The IM treatment group also demonstrated statistically significant pre- to posttest improvement accounting for 8.9% of the variance on an equally weighted multivariate reading composite measure (TOWRE and TSWRF). More impressive, however, was the posttest improvement accounting for 17% of the variance on a multivariate composite score that included the CTOPP tests Digit Naming, Letter Naming, Segmenting Nonwords, and Blending Nonwords and accounted for 9.9% of the variance on the CTOPP rapid automatized naming (RAN) test Letter Naming.

An alternative way to examine effect size is Hedges *g* (Howell, 2002). This statistic may be used to explain effect size as a percentage of growth, using a normal curve. Applying Hedge's *g* to the current results, the experimental group experienced a 20% growth on the CTOPP's RAN Letter Naming test and an 18% growth on the TOWER's Sight Word Efficiency. These growth rates compare favorably to the 15% growth identified in a meta-analysis of phonics instruction verses whole-word instruction conducted by the National Reading Panel's Committee on the Prevention of Reading Difficulties in Young Children (National Reading Panel, 2000).

The pre- to posttest reading achievement results suggest that improvements in timing and rhythmicity were associated with statistically significant improvements in three of the five major areas of measured reading: phonics, phonological awareness, and fluency. Yet, the results are not conclusive and must be moderated with a number of cautions. First, the experimental group did not demonstrate statistically significant increases on all the TOWRE's subtests. Second, although a significant improvement was observed on the CTOPP Letter Naming test, participants' performance on a similar test (Digit Naming) was not statistically significant. The key difference between the two tests is that the Letter Naming Test uses 26 letter stimuli, whereas the Digit Naming test's stimuli consist of 9 single-digit numbers. Third, on another measure of fluency (viz., WJ-III Reading Fluency) there was no statistically significant treatment effect. The lack of a significant effect for WJ-III Reading Fluency is at variance from a previous study involving high school students, wherein the experimental group demonstrated a statistically significant, 1-year grade level, improvement on the WJ-III Reading Fluency test (Taub, McGrew, & Lazarus, 2007).

Collectively, the current reading results suggest that students in the experimental IM treatment group demonstrated statistically significant improvements on more *fundamental* early reading skills (i.e., phonics and phonological awareness) and in their speed of processing basic lexical information (e.g., RAN for letters). However, with the exception of fluency of word recognition (i.e., Sight Word Efficiency test), students in the experimental group did not demonstrate statistically significant improvements at the single-word level.

Possible Causal Explanations: A Proposed Explanatory Framework and Preliminary Hypotheses

Previous IM intervention research reported statistically significant improvements in high schools students' performance on measures of reading recognition and reading fluency compared to a nontreatment control group (Taub, McGrew, & Lazarus, 2007). Similarly, IM-treated students with ADHD were reported to demonstrate statistically significant improvements in attention, reading, and language processing (Shaffer et al., 2001). This small collection of academically related studies, investigating direct reading achievement indicators and behaviors that exert an indirect causal influence on achievement (i.e., attention and concentration), are intriguing and suggest the need to focus efforts on understanding *why* improvements in timing and rhythmicity (via SMT interventions) display such far-point transfer effects.

In an effort to jump start efforts directed at understanding the underlying SMT–academic causal mechanisms, it is proposed that SMT-based research needs to be placed in a theoretically sound and empirically based research/conceptual framework. Furthermore, it is argued that the observed positive cross-domain or domain-general effect of SMT-based interventions result from improvements/changes within a domain-general cognitive mechanism (or a small number of domain-general mechanisms). Based on a review of relevant mental interval timekeeping literature, the following preliminary hypotheses are offered.

Master Internal Clock Based on Scalar Timing Theory

To deal with time, organisms (animal and human) have developed multiple timing systems that are active in more than 10 orders of magnitude with various degrees of precision (Buhusi & Meck, 2005). According to Buhusi and Meck, humans have developed three general classes of timing systems (circadian, interval, and millisecond timing), each associated with different behaviors and brain structures/mechanisms. The millisecond timing system, which is involved in a number of classes of human behavior (e.g., speech, music, motor control) and that primarily involves the brain structures of the cerebellum, basal ganglia, and the dorsolateral prefrontal cortex (Buhusi & Meck, 2005; Lewis & Miall, 2006), is most relevant for understanding SMT-based interventions.

Pacemaker-accumulator model. Human behavior based on the perception and timing in the range of seconds to minutes has traditionally been explained by the predominant model of interval timekeeping, namely, the *pacemaker–accumulator model* (PAM). The PAM, which is based on the *scalar expectancy or timing theory* (Church, 1984; Gibbon, Church, & Meck, 1984; Meck, 1983), "is relatively straightforward, and provides powerful explanations of both behavioral and physiological data" (Buhusi & Meck, 2005, p. 755).

Briefly, the PAM model implicates the processing of temporal information via three synchronized *modular information processing systems* (see Buhusi & Meck, 2005). The *clock* system consists of a dopaminergic *pacemaker* that regularly generates or emits neural ticks or pulses that are transferred (via a *gaiting* switch) to the *accumulator*, which accumulates ticks/pulses (neural counting) that correspond to a specific time interval. The raw representation of the stimulus duration in the accumulator is then transferred to working memory, a component of the PAM *memory* system. The contents of working memory are then compared against a *reference standard* in the long-term (reference) memory, the second component of the PAM memory system. Finally, the *decision* level of the PAM is conceptualized to consist of a *comparator* that determines an appropriate response based on a decision rule that involves a comparison between the interval duration value present in working memory and the corresponding duration value in reference memory. In other words, a comparison is made between the contents of reference memory (the standard) and working memory (viz., are they "close?").

Given evidence that supports a domain-general master internal clock central to many complex human behaviors (see Buhusi & Meck, 2005; Lewis & Miall, 2006), it is suggested that the *master internal clock* may be the mechanism that mediates SMT performance and intervention effects. It is hypothesized that SMT training improves human performance across a number of domains (e.g., reading and ADHD) via an increase in the *clock speed* of the master internal clock.

It is beyond the scope of the current study to describe the specific hypothesized brain mechanisms that produce a higher *clock speed* for the internal master clock. What is important to note in the current context is that mental interval timekeeping and temporal processing research has suggested that a *higher mental clock rate* enables individuals to perform specific sequences of mental operations faster and reduces the probability of interfering incidents (i.e., less disinhibition). These two conditions produce superior performance on cognitive tasks as well as more efficient basic information processing skills (Rammsayer & Brandler, in press).

The Master Mental Clock and Cognitive/Neuropsychological Constructs

The major components of PAM-based mental interval timekeeping have strong similarities to a number of domain-general cognitive mechanisms featured in contemporary cognitive information processing and/or neuropsychological research. Working memory, which is pivitol to PAM, is a central concept in major models of information processing. In addition, the PAM long-term (Buhusi & Meck, 2005) memory likely invokes early stages of memory consolidation in long-term memory or storage, another major component of information processing models of cognition. Furthermore, the *if-then* decision-making function of the PAM *comparator* is a function typically associated with skills involved with executive functioning (e.g., monitor, evaluate, change). Finally, research has implicated the important role of *attention* during the cognitively controlled portions of interval timing (Buhusi & Meck, 2005). Therefore, it is hypothesized that a conceptual cross-walk between the major components of the PAM master internal clock and contemporary cognitive information processing theories suggests that SMT performance and SMT transfer effects result in an increased efficiency in the functioning of the domain-general cognitive information processing mechanisms of (a) working memory, (b) executive functioning, and/or (c) controlled or executive attention.

Working Memory, Executive Functioning, and Executive Controlled Attention

Executive functioning (EF), which is also frequently called the *central executive system*, is a term used for a broad construct that represents a cluster of skills necessary for efficient and successful goal-directed behavior (Welsh, 2001). The EF constructs of planning, monitoring, inhibituation, and attention/concentration, elicit a range of basic cognitive processes (e.g., attention, perception, language, and memory) that are coordinated for a very specific purpose: subserving goal-directed behavior.

EF processes are believed to work in symphony to facilitate goal-directed task completion. Timing and processes related to mental timing are believed to be a component of executive function (Welsh, 2001), as is the utilization of executive functions during reading performance (Bull & Scerif, 2001). Because EF is an integration of a constellation of abilities necessary for the planning, self-monitoring/regulating, and evaluation of successful task completion, the area of self-regulated learning has received considerable attention with regard to a variety of cognitive activities (e.g., meta-cognition, pre-attentive processes, sluggish attentional shifting, specific strategy selection and implementation, inhibition, multitasking activities, task switching, maintenance of information under conditions of interference, and resistance to interference; Bull & Scerif, 2001; Borkowski, Carr, & Pressley, 1987; Kane, Bleckley, & Conway, 2001). The central role of EF in the enhancement of selective or controlled attention, the ability to switch between plans and strategies, and the inhibition of task-irrelevant information (intrusions) in working memory (Engle, Tuholski, Laughlin, & Conway, 1999; Passolunghi & Siegel, 2004) is consistent with theoretical and descriptive interpretations of SMT and interval time tracking models.

It is proposed that the *executive controlled attention model* of working memory (Engle, Kane, & Tuholski, 1999; Kane, Bleckley, Conway & Engle, 2001), which invokes the EF system, should be entertained as a potentially useful initial model to explain the domain-general effects of SMT-based interventions. Briefly, the executive controlled attention working memory model hypothesizes that individual differences in task performance are related to EF *controlled attention*. This means that individuals with higher working memory demonstrate better (or more efficient) use of attentional resources and are more able to resist interference during the encoding and retrieval processes than individuals with lower working memory. It is our hypothesis that SMT training does not improve working memory by increasing capacity, rather that SMT training may result in more *efficient* use of an individual's working memory system. The central role that the *general* capability to *efficiently process* information plays in task performance is consistent with a general mechanism explanation for the diversity of across-domain effects of SMT training. Central to the controlled attention working memory model is the role of EF. The alternative working memory view, which argues more for emphasis on underlying *modality-specific* working memory subprocesses (Palladino, Mammarella, & Vecchi, 2003), in contrast to resource-sharing models, presents a much more complex alternative model by which to explain positive SMT training effects across such diverse performance tasks (although it would be inappropriate to completely discard it as a possible explanation at this time). The search for a domain-general mechanism to explain SMT generalized training effects, such as the controlled attention working memory model, represents a more parsimonious approach that is believed to be preferred as formative attempts are made to describe and explain SMT training effects.

Finally, the recent suggestion that *g* or general intelligence (the most enduring and robust domain-general cognitive mechanism in the history of the psychometric study of intelligence) may be more a function of *temporal processing* and not necessarily reaction time (as measured by the traditional Hick paradigm; Rammsayer & Brandler, in press) suggests that mental interval timekeeping models (e.g, PAM) may describe and explain a primary elementary cognitive mechanism involved in most all complex human behavior. If *temporal g* exists, then the across-domain positive treatment effects of SMT training might be explained as the improvement of general neural efficiency via greater resolution of the temporal *g* internal clock.

Summary

This study investigated the effect of a SMT training intervention on elementary-school-age students' reading achievement. The observance of statistically significant improvements in the experimental group's performance on posttest measures of reading, when compared to the control group, is impressive given the nature of the *nonacademic* intervention. Yet, the results are not conclusive and are inconsistent in some cases. For example, the elementary school students scored significantly better on a timed single word recognition test, yet, there was no significant between-group difference on a measure that required reading short simple sentences (WJ-III Reading Fluency). Also, previous research with high school students reported a statistically significant relationship between SMT improvements and reading fluency. One possible explanation for the divergent developmental intervention effect findings is that elementary school students are *learning how to read*, whereas high school students are *reading to learn*. In other words, high school students have mastered or automatized their reading skills, whereas the elementary school students are learning how to read.

Nevertheless, the automatization of critical early reading skills (viz., phonics, phonological awareness skills, and RAN performance), which emerge primarily during the early school grades, are the specific areas where the elementary-aged experimental participants demonstrated the most significant improvements in the current study. It is also possible that studies (the current study, inclusive) that have reported improvements in timing and rhythmicity over short periods (3–4 weeks) may only demonstrate significant effects on the processing of overlearned (automatized) information, in contrast to the more deliberate or controlled learning of new information. This may also explain why golfers, who presumably have overlearned their golf swing, become more accurate with improvements in timing/rhythmicity.

It is believed that subcomponents of the constellation of executive functioning are effected by SMT interventions. Because of the cross-domain influence of working memory on task completion, the executive controlled attention model of working memory, which is heavily dependent on the executive functioning system, was hypothesized as a potentially useful model for conceptualizing SMT research and for interpreting research findings. The executive controlled aspect of working memory was suggested as a possible general cognitive mechanism responsible for the observed positive influence of SMT training across such diverse domains as academics, athletics, and attention/concentration.

Limitations and Future Research

This study may be limited by participants' parents self-selection to have their child attend a public charter school receiving Title 1 funding. Participants may also have been more similar on several demographic variables (e.g., ethnicity, socioeconomic status) than would be found in public school settings.

Because of the relatively small sample size it was not possible to make a distinction between students receiving special education services and those who were not. It is recommended that future studies examine this difference as well as investigate differential SMT training effects with regular education students experiencing academic difficulties. It is also recommended that future studies investigate SMT training effects with students who were unable to graduate or progress to the next grade level because they did not reach a threshold score on high-stakes tests of academic achievement.

Finally, in the present study posttests were administered immediately after SMT training; therefore the stability of the observed positive effects of SMT training on the academic achievement dependent variables is not known. It is recommended that future studies investigate the consistency of the observed positive effects of SMT training on academic achievement over an extended period.

References

Barkley, R. (1997). ADHD and the nature of self control. New York: Guilford.

Borkowski, J.G., Carr, M., & Pressley, M. (1987). "Spontaneous" strategy use: Perspectives from metacognitive theory. Intelligence, 11, 61–75.

Brown, S.W., & Bennett, E.D. (2006). The role of practice and automaticity in temporal and nontemporal dual-task performance. Psychological Research, In Press.

Buhusi, C.V. & Meck, W.H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. Neuroscience, 6, 755–765.

Bull, R., & Scerif, G. (2001). Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. Developmental Neuropsychology, 19, 273–293.

Carroll, J.B. (1993). Human cognitive abilities: A survey of factor analytic studies. New York: Cambridge University Press.

Church, R.M. (1984). Properties of the internal clock. Annals of the New York Academy of Sciences, 433, 566–582.

Daly, E.J., III, & McCurdy, M. (2002). Getting it right so they can get it right: An overview of the special series. School Psychology Review, 31, 453–458.

Engle, R.W., Kane, M.J., & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control (pp. 102–134). New York: Cambridge University Press.

Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. Journal of Experimental Psychology: General, 128, 309–331.

Gibbon, J., Church, R.M., & Meck, W.H. (1984). Scalar timing in memory. Annals of the New York Academy of Sciences, 423, 52–77.

Howell, D.C. (2002). Statistical methods for psychology (5th ed.). Pacific Grove, CA: Duxbury.

Individuals with Disabilities Education Improvement Act, 20 U.S.C. § 1400 et seq. (2004).

Kane, M.J., Bleckley, M.K., Conway, A.R.A., & Engle, R.W. (2001). A controlled-attention view of working-memory capacity. Journal of Experimental Psychology General, 130, 169–183.

Katzir, T., & Paré-Blagoev, J. (2006). Applying cognitive neuroscience research to education: The case of literacy. Educational Psychologist, 41, 53–74.

Keith, T.Z. (2006). Multiple regression and beyond. Boston, MA: Pearson Education.

Lewis, P.A., & Miall, R.C. (2006). A right hemispheric prefrontal system for cognitive time measurement. Behavioural Processes, 71, 226–234.

Libkuman, T. M., & Otani, H. (2002). Training in timing improves accuracy in golf. Journal of General Psychology, 129, 77–96.

Mather, N., Hamill, D., Allen, E.A., & Roberts, R. (2004). Test of Silent Word Reading Fluency (TOSWRF). Austin, TX: Pro-Ed.

Mather, N., Wedling, B. J., & Woodcock, R. W. (2001). Essentials of WJ-III Tests of Achievement Assessment. New York: John Wiley & Sons, Inc.

McGee, R. Brodeur, D., Symons, D., & Fahie, C. (2004). Time perception: Does it distinguish ADHD and RD children in a clinical sample? Journal of Abnormal Child Psychology, 32, 481–490.

McGrew, K.S., & Woodcock, R.W. (2001). Technical manual. Woodcock-Johnson III. Itasca, IL: Riverside Publishing.

Meck, W.H. (1983). Selective adjustment to the speed of internal clock and memory processes. Journal of Experimental Psychology: Animal Behavior Processes, 9, 171–201.

National Reading Panel (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Washington, DC: National Institute of Child Health and Human Development.

Palladino, P., Mammarella, N., & Vecchi, T. (2003). Modality-specific effects in inhibitory mechanisms: The interaction of peripheral and central components in working memory. Brain and Cognition, 53, 263–267.

Passolunghi, M.C., & Siegel, L.S. (2004). Working memory and access to numerical information in children with disability in mathematics. Journal of Experimental Child Psychology, 88, 348–367.

Rammsayer, T.H., Brandler, S. (2007). Performance on temporal information processing as an index of general intelligence. Intelligence, 35, 123–139.

Shaffer, R.J., Jacokes, L.E., Cassily, J.F., Greenspan, R.F., Tuchman, P.J., & Stemmer, P.J. (2001). Effect of interactive metronome training on children with ADHD. American Journal of Occupational Therapy, 55, 155–161.

Shaywitz, S.E., Fletcher, J.M., Holahan, J.M., Shneider, A.E., Marchione, K.E., Stuebing, K.K., et al. (1999). Persistence of dyslexia: The Connecticut longitudinal study at adolescence. Pediatrics, 104, 1351–1359.

Shaywitz, S.E., & Shaywitz, B.A. (2005). Dyslexia (specific reading disability). Biological Psychiatry, 57, 1301–1309.

Shaywitz, S.E., Shaywitz, B.A., Fulbright, R.K., Skudlarski, P., Mencl, W.E., Constable, R.T., et al. (2003). Neural systems for compensation and persistence: Young adult outcome of childhood reading disability. Biological Psychiatry, 54, 25–33.

Sheridan, S. (2004). Editor's comments: Contemporary research on curriculum-based measurement. School Psychology Review, 33, 187.

Stankov, L., Horn, J.L., & Roy, T. (1980). On the relationship between Gf/Gc theory and Jensen's Level I/Level II theory. Journal of Educational Psychology, 72, 796–809.

Suss, H.-M., Oberauer, K., Wittmann, W.W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability—and a little bit more. Intelligence, 30, 261–288.

Tallal, P., Miller, S., Jenkins, B., & Merzenich, M. (1997). The role of temporal processing in developmental language-based learning disorders: Research and clinical implications. In B. Blachman (Ed.), Foundations of reading acquisition (pp. 21–47). Hillsdale, NJ: Erlbaum.

Taub, G.E., McGrew, K.S., & Lazarus, P.J. (2007). The effects of training in timing and rhythm on reading achievement. Manuscript submitted for publication.

Thompson, B. (1999). Improving research clarity and usefulness with effect size indices as supplements to statistical significance tests. Exceptional Children, 65, 329–337.

Torgesen, J.K., Wagner, R.K., & Rashotte, C.A. (1999). Test of Word Reading Efficiency (TOWRE). Austin, TX: Pro-Ed.

U.S. Department of Education (2002). No Child Left Behind: A desktop reference. Washington, DC: Author.

Vorberg, D., & Fuchs, A. F. (2004). Tapping along perturbed metronomes: Automatic phase responses to complex perturbation patterns but no long-term effects. Paper presented at the 9th Rhythm Perception and Production Workshop. Abstract retrieved April 12, 2004 from http://www.rppw.org/2004.

Wagner, R.K., Torgesen, J.K., & Rashotte, C.A. (1999). Comprehensive Test of Phonological Processing (CTOPP). Austin, TX: Pro-Ed.

Waber, D.P., Marcus, D.J., Forbes, P.W., Bellinger, D.C., Weiler, M.D., Sorensen, L.G., et al. (2003). Motor sequence learning and reading ability: Is poor reading associated with sequencing deficits? Journal of Experimental Child Psychology, 84, 338–354.

Welsh, M.C. (2001). Developmental and clinical variations in executive functions. In B.L. Molfese & V.J. Molfese (Eds.), Developmental variations in learning: Applications to social, executive function, language and reading skills (pp. 139–185). Mahwah, NJ: Erlbaum.

Woodcock, R.W., Mather, N., & McGrew, K.S. (2001). Woodcock Johnson Psychoeducational Battery, 3rd ed. Chicago: Riverside.

Wolff, P.H. (2002). Timing precision and rhythm in developmental dyslexia. Reading and Writing, 15, 179–206.

Zachopoulou, E., Mantis, K., & Psalti, M. (in press). The effect of rhythmic ability on the forehand in tennis. Psychological Research.

■ REVIEW

How Do We Tell Time?

DEAN V. BUONOMANO and UMA R. KARMARKAR
Departments of Neurobiology and Psychology
and Brain Research Institute
University of California, Los Angeles

Animals time events on scales that range more than 10 orders of magnitude—from microseconds to days. This review focuses on timing that occurs in the range of tens to hundreds of milliseconds. It is within this range that virtually all the temporal cues for speech discrimination, and haptic and visual processing, occur. Additionally, on the motor side, it is on this scale that timing of fine motor movements takes place. To date, psychophysical data indicate that for many tasks there is a centralized timing mechanism, but that there are separate networks for different intervals. These data are supported by experiments that show that training to discriminate between two intervals generalizes to different modalities, but not different intervals. The mechanistic underpinnings of timing are not known. However various models have been proposed, they can be divided into labeled-line models and population clocks. In labeled-line models, different intervals are coded by activity in independent and discrete populations of neurons. In population models, time is coded by the population activity of a large group of neurons, and timing requires dynamic interaction between neurons. Population models are generally better suited for parallel processing of interval, duration, order, and sequence cues and are thus more likely to underlie timing in the range of tens to hundreds of milliseconds.
NEUROSCIENTIST 8(1):42–51, 2002

KEY WORDS *Timing, Temporal processing, Models, Interval discrimination, Learning, Neural networks*

> Temporal integration is not found exclusively in language; the coordination of leg movements in insects, the song of birds, the control of trotting and pacing in a gaited horse, the rat running the maze, the architect designing a house, and the carpenter sawing a board, present a problem of sequences of action which cannot be explained in terms of succession of external stimuli.

This quote is from an article titled “The Problem of Serial Order in Behavior” by Karl Lashley (1951/1960). Lashley wrote the article because he felt that temporal processing was “the most important and also the most neglected problem of cerebral physiology.” The article was written 2 years after Donald Hebb wrote the *Organization of Behavior*, the book in which Hebb presented his influential theory on the rules that govern synaptic plasticity. However, in contrast to the topic addressed by Hebb, the topic discussed by Lashley has not seen significant advances in the past half century.

A fundamental part of sensory processing is pattern recognition, that is, how central neurons develop selective responses to the spatial and temporal patterns of activity coming from primary sensory neurons. We can decompose sensory stimuli into spatial and temporal components. Spatial stimuli refer to those that can be discriminated based on a static “snapshot” of which neurons are active, that is, the spatial arrangement of active neurons. Discriminating the orientation of bars of light, or letters of the alphabet, falls into this category. In the past 50 years, much progress has been made on this front. Indeed, the fields of synaptic plasticity and self-organizing topographic maps explain how neurons can develop responses to simple spatial stimuli (for reviews, see Anderson and Rosenfeld 1988; Buonomano and Merzenich 1998). These advances, however, say very little about how neurons develop selective responses to temporal patterns. Temporal patterns refer to those in which the order, duration, or interval between the activation of sensory neurons is required for stimulus discrimination. The duration of flashed bars of light and the voice-onset time of phonemes are examples of temporal stimuli. Without an understanding of the neural mechanisms underlying temporal processing, it will not be possible to understand how the brain processes complex real-world stimuli, which are characterized by both their spatial and temporal features. For example, speech recognition, one of the most complex forms of pattern recognition, relies on both spatial and temporal processing (Tallal 1994). Indeed, one of the difficulties in understanding how the brain processes speech, and in the construction of artificial systems capable of speech recognition, stems from underestimating the importance of temporal information in speech (Shannon and others 1995). In addition to this and other forms of sensory processing, timing plays a fundamental role in motor coordination. Given

I thank Jason Goldsmith for comments on earlier versions of this manuscript. This research was supported by the Alfred P. Sloan and EJLB Foundations and by the NSF.

Address correspondence to: Dean V. Buonomano, Brain Research Institute, University of California, Los Angeles, Box 951761, Los Angeles, CA. 90095 (e-mail: dbuono@ucla.edu).

ISSN 1073-8584

the inherent time-varying nature of our environment and our interaction with it, it is fundamental to understand the neural basis of how the brain processes time.

Temporal Processing: Time Scales

The terms *temporal processing*, *temporal integration*, and *timing* are used to describe a wide range of different phenomena, which often results in ambiguity in the literature. One source of ambiguity is the large scale over which animals process temporal information or generate timed behaviors. All together the brain processes temporal information over a range of at least 10 orders of magnitude from microseconds to daily circadian rhythms—and the above terms are used to refer to all of them. Based on the relevant time scales and the supposed underlying neural mechanisms, we can categorize timing into four different time scales: microseconds, milliseconds, seconds, and circadian rhythms (Fig. 1). These classes are not meant to represent discrete nonoverlapping types of processing. Instead, they represent a simplified division of the number of ranges of temporal processing that rely on different neural mechanisms.

Microseconds

Microsecond temporal processing is used primarily for the detection of interaural delays, the detection of electric fields in electric fish, and echo-location in bats (in which the relevant delays extend up to 10 msec). The best understood system is that used for sound localization. In humans it takes sound approximately 600 to 700 μs to travel the distance between the left and right ear. The auditory system uses these intervals to calculate the spatial location of the sound source. A relatively simple but extremely sensitive mechanism is used to determine the microsecond intervals for sound localization. A sound arriving in each ear will activate neurons in the cochlear nucleus. The axons from these neurons function as delay lines; that is, the distance a action potential has to travel is proportional to the time it takes. Neurons in the medial superior olive function as coincidence detectors and use the delays to respond selectively to different intervals. Together these neurons establish a topographic map of auditory space (Carr 1993).

Milliseconds

Millisecond processing will be defined as that above 10 msec and below 500 to 1000 msec. Sensory processing within this range is often referred to as perceptual timing: "below 0.5 sec information processing is of a highly perceptual nature, fast parallel and not accessible to cognitive control" (Michon 1985, p. 21). Millisecond processing is perhaps the most sophisticated and the least well understood. Virtually all the temporal cues for speech and vocalization discrimination, and many of the cues in music perception, fall within this range. Additionally, much of the motion processing in the visual and somatosensory system occurs on this scale. On the motor side, it is within the range of tens to hundreds of milliseconds that fine motor coordination operates in. Thus, the ability of athletes and musicians to perform extraordinary physical feats relies on sophisticated neural mechanisms capable of producing well-timed and orchestrated events in the millisecond range.

Seconds

Timing on scales longer than a second are often referred to as time estimation and thought to rely on conscious and cognitive control (Rammsayer and Lima 1991). Millisecond and second processing are thought to rely on different mechanisms based on psychophysical and pharmacological experiments. Rammsayer and Lima showed that interval discrimination of 50 msec intervals was unaffected by cognitive load, whereas intervals of 1 sec were. Additionally, pharmacological manipulations can differentially affect millisecond and second processing (see below). In addition to time estimation, there are various behaviors that rely on pattern generators operating in this time scale—such as breathing and locomotion. For reviews on timing in the range of seconds and minutes, see Gibbon and others (1997) and Matell and Meck (2000).

Circadian Rhythms

Animals also track time through daily circadian rhythms. In addition to the daily sleep-wake cycles, regulation of hormone levels, thermoregulation, and appetite cycles are occurring on the scale of hours and days. Sleep-wake cycles are a good example of a behavior controlled by an internal clock. Physiological measures in both plants and animals can be shown to exhibit an approximately 24-h rhythm, even in the absence of external stimuli. The clock controlling circadian rhythms is not immutable; its phase can be shifted and entrained by external cues. Studies in various organisms, including *Drosophila* and mice, have revealed that circadian clocks are composed of molecular/biochemical pathways regulating transcription and translation in autoregulatory feedback loops (for a review of the molecular mechanisms of circadian clocks, see King and Takahashi 2000).

In the current review, focus will be on time perception temporal processing occurring in the range of tens to hundreds of milliseconds. This time scale is fundamental to sensory processing in the auditory, visual, and somatosensory modalities. As mentioned above, motor coordination and speech perception exemplify how sophisticated temporal processing can be on the millisecond scale. During continuous speech, syllables are generated every 200 to 400 msec. The sequential arrangement of syllables is important in speech recognition (e.g., "la-dy" vs. "de-lay"). Similarly, the duration of each syllable is critical, as is the interval between syllables (e.g., by emphasizing the timing of Jimi Hendrix's famous mondegreen "kiss the sky," it is easier to distinguish it from "kiss this guy"). Additionally, the temporal structure within each syllable and phoneme also contributes to discrimination. For example, the voice-onset

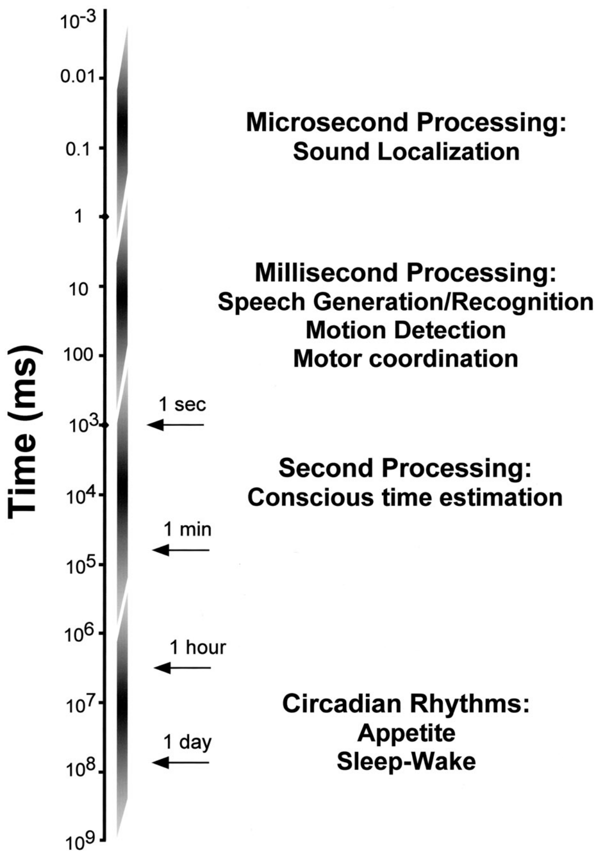


Fig. 1. Scales of temporal processing. Humans process temporal information over a scale of at least 10 orders of magnitude. On one extreme, we detect the delay required for sound to travel from one ear to the other. These delays are on the order of tens to hundreds of microseconds. On the other extreme, we exhibit daily physiological oscillations, such as our sleep-wake cycle. These circadian rhythms are controlled by molecular/biochemical oscillators. Temporal processing on the scale of tens and hundreds of milliseconds is probably the most sophisticated and complex and is fundamental for speech processing. Time estimation refers to processing in the range of seconds and minutes and is generally seen as the conscious perception of time.

time (the time between air release and vocal cord vibration) and transition duration of formants are used for the discrimination of individual consonant-vowel syllables (Tallal 1994). Prosodic cues such as pauses and duration of speech segments are used to determine semantic content (Lehiste and others 1976).

Temporal processing on the scale of microseconds, seconds, and days seems to be less complex than millisecond processing. For example, microsecond processing for interaural delay detection is not capable of duration or sequence discrimination. Timing in the range of seconds and minutes generally involves conscious estimation of intervals and is not used for sequence or parallel processing of multiple temporal cues or of periodic pattern generation. Circadian rhythms are likely to be controlled by biological clocks and exhibit less flexibility than temporal processing on the shorter time scales. For example, the internal clock controlling circadian rhythms cannot be instantly reset (thus jet lag). In contrast, time perception and time estimation can begin at the onset of any stimulus. Processing on the millisecond range seems to be the most complex. In speech we are processing the temporal structure of phonemes, the prosody of speech, and sequence of speech segments all in parallel. Additionally, temporal discrimination can exhibit a higher-order form of processing referred to as temporal invariance: we can identify the same speech segments or tone sequences at a range of speeds, as long as the ratios between different events are similar. Thus, the neural mechanisms underlying temporal processing in the millisecond range are likely to be complex and may or may not rely on independent mechanisms to solve specific components of temporal processing, such as order, duration, intervals, inter- and intramodality timing, and motor timing.

Central versus Distributed Mechanisms

A fundamental question regarding temporal processing is whether it relies on a single centralized mechanism or is distributed throughout different areas. If timing is centralized, then an interval discrimination task in the somatosensory, visual, or auditory modality would use the same group of neurons. Additionally, motor tasks requiring carefully timed responses would also rely on the same system. In this view, timing in the nervous system would be analogous to that in computers, in which a central clock sends out information to many other components of the computer. In contrast, in a distributed system various regions of the brain would process time, and the locations used would depend on the modality and task at hand. Thus, different parts of the brain would be involved in timing in somatosensory, auditory, visual, or motor tasks.

In the psychological literature on timing, by far the most influential model has been the internal clock model (Creelman 1962; Treisman 1963). Internal clocks are hypothetical mechanisms in which a neural pacemaker generates pulses; the number of pulses relating to a physical time interval is recorded by a counter. Internal clock models are generally centralized: one clock is used for all timing tasks.

Centralized and distributed mechanisms can be subdivided into models in which the same neurons are timing all intervals or models in which different neurons time different intervals. For example, we can use the same watch to time both 100 or a 500 msec intervals. However, one could imagine a system in which the initial event triggered an array of watches, each one devoted to a fixed interval: 100, 200, . . ., 500 msec. In this review, the former model will be referred to as a clock model and the latter as a labeled line or an interval-based model (Ivry 1996).

Correlations between Temporal Tasks

The majority of the timing studies in humans rely on interval discrimination tasks (Fig. 2*A*). In a typical task, two brief tones separated by a standard interval (e.g., 100 msec) or a comparison interval (standard + ΔT) are presented to the subject. The order of the presentation of the standard and comparison intervals is randomized. The subject is required to make a judgment as to whether the longer interval was the first or second. Depending on the task design, the difference in milliseconds between the short and long interval (ΔT) is adaptively changed according to performance, which allows the calculation of an interval discrimination threshold (Wright and others 1997).

If timing relies on a centralized mechanism, a correlation between different timing tasks would be expected. That is, are individuals that are good at discriminating auditory intervals also good at discriminating somatosensory intervals? Two types of correlations can be analyzed, those between different modalities for the same interval and between different intervals in the same modality. High correlations in the former analysis would suggest a central timekeeping mechanism that is used in all modalities, but there could be independent timing mechanisms for each interval. In the latter, if a high correlation is observed between intervals, the analysis would support the notion that one central clock is being used for all intervals.

A study by Keele and others (1985) examined the correlations between a motor task and an auditory interval discrimination task. Moderate correlations (R^2 of approximately 0.5) between tapping and tone discrimination using target intervals of 400 msec were observed. A second study (Spencer and others 2000) also reports moderate correlations between both a 400 msec tapping and tone task ($R^2 = 0.39$) and an 800 msec target interval ($R^2 = 0.36$). This study also revealed a correlation between the 400 and 800 msec perception task ($R^2 = 0.54$). Figure 2*B* shows plots of the correlations between different conditions in an auditory discrimination task (Karmarkar and Buonomano, unpublished data). Four conditions were examined: 50 msec–1 kHz, 50 msec–4 kHz, 100 msec–1 kHz, and 200 msec–1 kHz. The results show significant correlations between 50 msec–1 kHz and 50 msec–4 kHz, as well as between 50 msec–1 kHz and 100 msec–1 kHz, but not between 50 msec–1 kHz and 200 msec–1 kHz. Together these results favor a centralized timing mechanism shared by sensory and motor systems for similar intervals. However, the lack of correlation between the 50 msec–1 kHz and 200 msec–1 kHz suggests that there may be distinct mechanisms for 50 msec and 200 msec timing. It should be stressed that the results from correlations studies are suggestive, in that they could also be attributed to experience-dependent generalization, rather than common underlying mechanisms.

Intermodal Timing

Data from some interval discrimination tasks support the notion of distributed timing. Specifically, some studies have examined tasks in which intervals are bounded by intermodal stimuli. Interval discrimination of intervals bounded by a tone and a flash of light (or a flash and a

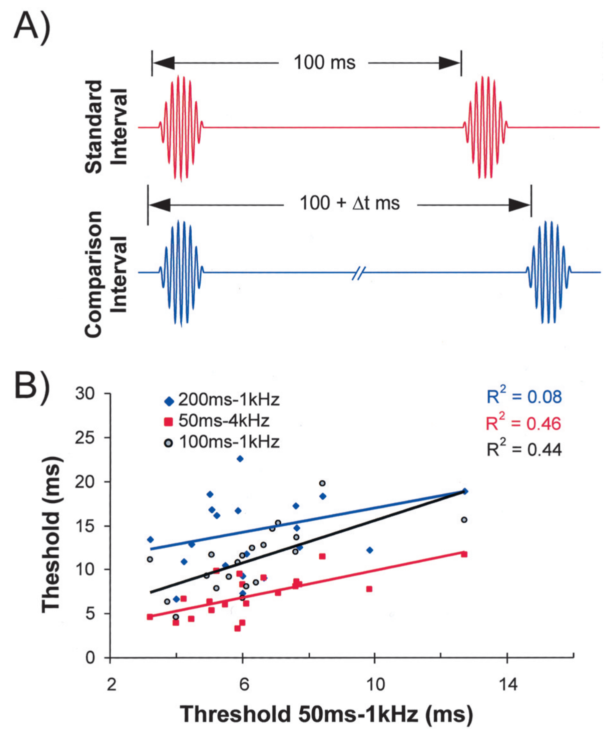


Fig. 2. Intrasubject correlations between interval discrimination tasks. *A*. Interval discrimination. Interval tasks can be designed in various ways. In one design, a standard and comparison interval are presented in random order, the subject has to decide whether the comparison interval (the longer one) came first or second. Both intervals are bounded by two brief tones of a fixed frequency. The standard interval is always the same length, whereas the comparison interval is equal to the standard + ΔT, where ΔT changes according to performance. Different task conditions are examined by varying the standard intervals and the frequency of the tones. *B*. Intrasubject correlations between different interval discrimination task conditions. Performance is well correlated for the same interval at different frequencies ($R^2 = 0.46$, $P < 0.005$). There is also a significant correlation between the 50 × 100 msec intervals ($R^2 = 0.44$, $P < 0.005$), but not between the 50 and 200 msec intervals ($R^2 = 0.08$, $P = 0.22$).

tone) is significantly worse than intervals bounded by two tones or two flashes (Rosseau and others 1983; Grondin and Rousseau 1991). Interestingly, intermodality discrimination is impaired relative to intramodality timing for subsecond processing, but not for 1 sec intervals (Rosseau and others 1983). Not only is intermodality discrimination less accurate than intramodality discrimination, but even within a given modality, discrimination is impaired by using intervals bounded by different stimulus characteristics (Divenyi and Danner 1977; Grondin and Rousseau 1991). Thus, interval discrimination of a 250 msec interval marked by two 1 kHz tones is better than the same intervals marked by a 1 kHz tone and a noise burst (Grondin and Rousseau 1991).

These data can be used to argue for distributed timing because a centralized timer may be expected to time events arriving through different channels as well as events arriving through the same channel. However, an alternative explanation is that intermodal timing is simply a more difficult task because it requires a shift of attention from one modality to the other.

Anatomical Localization

If timing is centralized, it is important to ask: where is it located? Various structures have been implicated in timing. One such area is the right parietal cortex. A recent study showed that stroke patients with right hemisphere parietal lesions, but not left hemisphere lesions, exhibit a selective deficit for 300 and 600 msec interval discrimination (Harrington, Haaland, and Knight 1998). A second structure implicated in timing is the basal ganglia, although it is generally thought to contribute to timing on the scale of seconds. Two studies have shown that Parkinson patients exhibit deficits in temporal discrimination in the millisecond range, but not in frequency discrimination (Artieda and others 1992; Harrington and others 1998). These data are indirect because it is possible that Parkinson's effect on timing is due to secondary effects on structures other than the basal ganglia. In addition to the cortex and basal ganglia, the cerebellum has also been proposed to underlie timing. Because it is the structure that has been the best studied in relation to temporal processing, it will be discussed in detail below.

Cerebellum

Braitenberg (1967) suggested that one of the main functions of the cerebellum was timing. He made the specific proposal that the axons of the granule cells (parallel fibers) functioned as delay lines. This hypothesis is currently not accepted, primarily because given the conduction velocity of parallel fibers, it would require a 5-cm-long fiber to create a 100 msec delay. Furthermore, because granule cells are not excitatory, nor are there excitatory loops in the cerebellum, the cerebellar architecture does not support "excitatory chains" to implement longer delays.

Although the mechanisms are debated (see below), there is growing experimental support for a cerebellar role in timing (for a review, see Ivry 1996). This is particularly true for motor timing. One of the best studied systems regarding the timing of motor responses is eyeblink conditioning. In this form of conditioning, an animal receives paired presentation of a tone and a puff of air to the cornea. As a result of this training, animals learn to blink in response to the tone alone. Animals do not learn to blink arbitrarily on hearing the tone, but blink at a time equal to the interval between the tone and air puff presented during training. Lesions to the cerebellar cortex abolish the timing of the conditioned response, without eliminating it (Perrett and others 1993).

There is also support for a role of the cerebellum in forms of sensory timing, such as interval discrimination. Ivry and Keele (1989) showed that subjects with cerebellar lesions were less accurate in a 400 msec interval discrimination as compared with control subjects with cortical lesions. Other studies have shown deficits in the discrimination of phonemes differing in their temporal structure in subjects with bilateral cerebellar lesions (Ackermann and others 1997). Imaging studies have shown that the cerebellar vermis is activated during a 300 msec interval discrimination task (Jueptner and others 1995). However, in this study, the control task did not require decision making or stimulus comparisons, and other areas such as the right thalamus and basal ganglia were also active. Furthermore, it has been suggested that the observed increases in blood flow may reflect cerebellar involvement in complex stimulus analysis and not necessarily an explicit role in timing (Ackermann and others 1999).

Various lines of evidence suggest that one or more structures may play a predominant role in timing and function as a central time-keeping structure. However, to date no study has shown that a given lesion or disease eliminates temporal processing. This could be taken as indirect evidence for distributed timing mechanisms, in that none of the lesion studies produce a global multimodal sensory-motor breakdown in timing.

Interval Discrimination Learning

One question that has not been examined carefully until recently is whether interval discrimination undergoes perceptual learning. That is, does temporal resolution increase with practice. One of the first studies to examine this issue reported no perceptual learning (Rammsayer 1994). In this study, subjects were trained on 50 msec intervals for 10 min a day for 4 weeks. More recent studies have all reported improvement of interval discrimination with practice (Wright and others 1997; Nagarajan and others 1998; Westheimer 1999). In these studies, subjects were generally trained for an hour a day for 10 days.

Generalization of Interval Discrimination

In addition to showing that the neural mechanisms underlying timing can be fine tuned with experience, learning studies provide a means to examine the issue of central versus distributed timing. Specifically, we can ask if after training on a 100 msec interval bounded by 1 kHz tones the performance improves for different intervals and frequencies. If there is a single central timer that relies on a clock mechanism, generalization to both different intervals and different marker conditions should be observed.

The first study to address this issue revealed that after training on 100 msec intervals marked by 1 kHz tones, subjects showed complete generalization to the same interval marked by 4 kHz tones (Fig. 3) (Wright and others 1997). Subsequent work revealed that intermodal generalization was observed (Nagarajan and others 1998). Training on a somatosensory interval discrimination task resulted in improvement on an auditory task for the same intervals. Both studies revealed little or no generalization to novel intervals presented with the same markers as the trained condition. That is, despite improve-

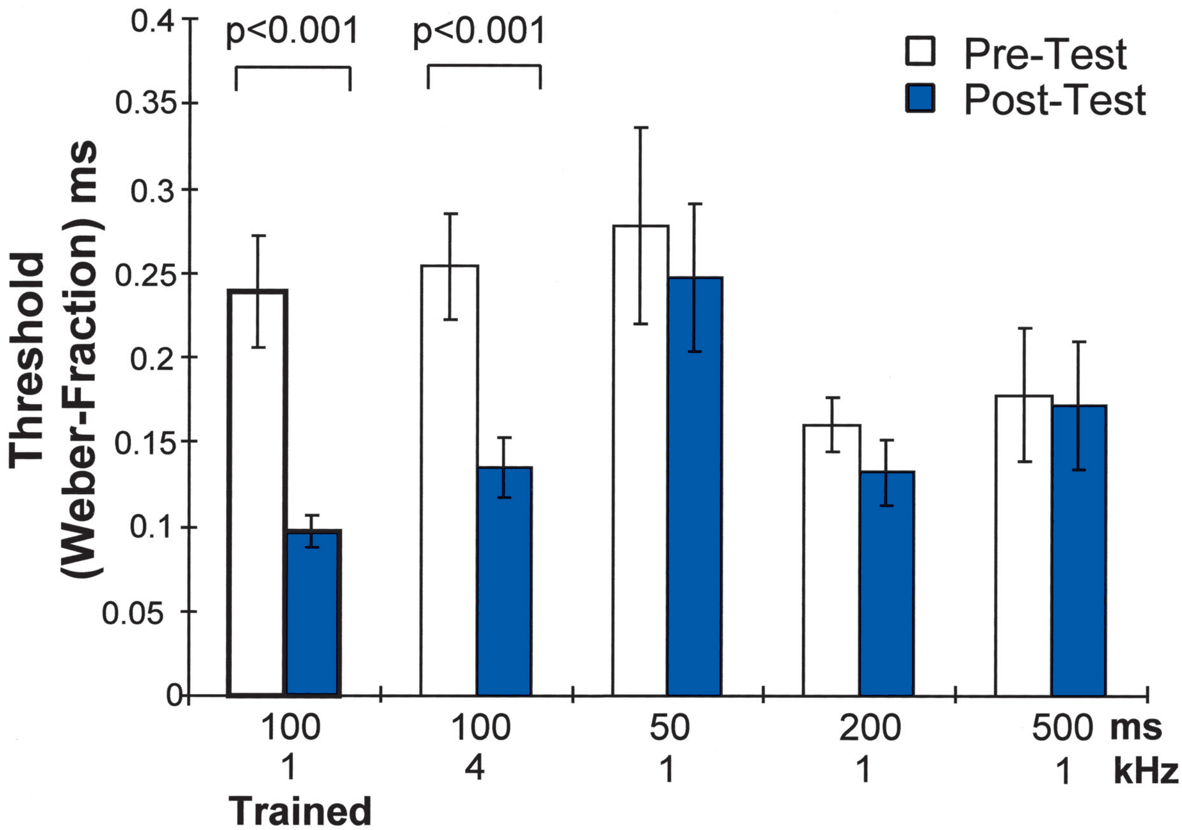


Fig. 3. Interval discrimination learning generalized across frequencies but not intervals. Subjects were trained on the 100 ms–1 kHz conditions for 10 days. The pre- and posttest thresholds revealed significant differences only for the trained condition, and the 100 ms–4 kHz condition. Modified from Wright and others (1997).

ment on the trained 100 msec interval, there was no improvement on 50 or 200 msec intervals. Together, these studies show that interval learning does not generalize in the temporal domain (different intervals) but does generalize in the spatial domain (different markers). This conclusion is also supported by results in the visual modality. Westheimer (1999) reported that training on a 500 msec duration visual stimulus presented to one visual hemifield generalized to the other hemifield. Even more surprising, training on an auditory task appears to result in an interval-specific improvement in a motor task requiring that the subjects tap their fingers to mark specific intervals (Meegan and others 2000).

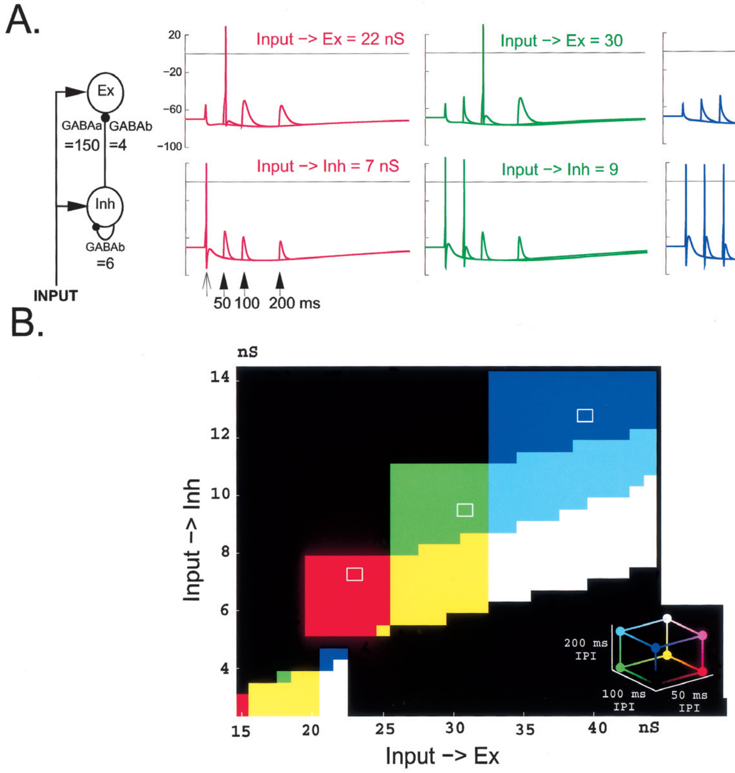
The simplest interpretation of these data is that there is a centralized clock for each interval, because the improvement is interval specific but generalized across modalities (somatosensory to auditory, and auditory to motor). The caveat in this interpretation is that it is possible that in these tasks learning occurs as a result of interval-specific cognitive processes other than temporal processing per se. For example, because interval discrimination requires comparing the test interval to a standard interval, improvement could rely on better representation or memory of the standard interval. Such an explanation would be consistent with the generalization across different stimulus markers and modalities, as well as the lack of generalization to novel intervals.

Psychopharmacology of Timing

Psychopharmacological experiments have also been used to probe the mechanisms underlying timing and to determine whether different time scales of processing rely on different neural systems. Numerous drugs have subjectively been reported to alter time estimation, that is, temporal processing in the seconds and minutes range, but few drug studies have carefully examined timing. One well-established finding is that dopamine antagonists produce temporal overshoot ("slowing of the clock"), and stimulants such as methamphetamine produce temporal undershoots ("speeding up the clock"; for a review, see Meck 1996). Few studies have examined pharmacological effects on temporal processing below a second. Rammsayer (1999) showed in human psychophysical experiments that the dopaminergic antagonist, haloperidol, significantly impaired discrimination thresholds for 100 msec and 1 sec intervals. Remoxipride, a dopamine antagonist that is more selective for D2 receptors, impaired processing on the scale of a second but not for 50 msec intervals (Rammsayer 1997). Experiments with benzodiazepines also support the dissociation between millisecond and second processing, by showing that performance in a 50 or 100 msec task is unaffected, whereas performance on a 1 sec task is made significantly worse (Rammsayer 1999, 1992). Together these results show that two distinct drug

Box 1: Interval Selectivity in Disynaptic Circuits

Computer simulations show how disynaptic circuits can exhibit interval selectivity. The circuit is composed of a single excitatory (Ex) and inhibitory (Inh) neuron, and there are five synapses: Input→Ex, Input→Inh, Inh_{fast}→Ex, Inh_{slow}→Ex, Inh_{slow}→Inh. The excitatory synapses exhibit paired-pulse facilitation (PPF), the inhibitory neuron produces both a fast ($GABA_A$) and a slow ($GABA_B$) inhibitory postsynaptic potential (IPSP) on the Ex neuron. Part *A* shows traces from the Ex and Inh cells for three different sets of synaptic strengths (red, green, and blue). Each graph shows the overlaid responses to three different intervals. By changing the strengths of the Input→Ex and Input→Inh connections in parallel, it is possible to tune the Ex unit to respond selectively to either 50, 100, or 200 msec intervals. With relatively weak inputs to both the Ex and Inh cell (red traces), the first pulse generates a supra- and subthreshold response in the Inh and Ex units, respectively. At 50 msec, the second pulse is suprathreshold in the Ex unit (even though it is riding a slow IPSP elicited by the first spike in the Inh unit), owing to PPF, which peaks at 50 msec. The second pulse, at any interval, does not generate a fast IPSP because the Inh unit did not fire owing to the $GABA_B$-mediated slow IPSP. If the strength of both inputs is increased (green traces), the Ex unit fires exclusively to the 100 msec pulse. It no longer fires to the 50 msec pulse because as a result of the increased input, the Inh unit fires in response to the second pulse at 50 msec producing a fast IPSP in the Ex unit, which prevents it from firing. If we continue to increase the strength of both inputs (blue traces), through a similar mechanism, the Ex unit fires exclusively to the 200 msec interpulse interval (IPI). Part *B* displays a parametric analysis of the interval selectivity described above in synapse space. The strength of the Input→Ex and Input→Inh was parametrically varied over a range of weights. The results are represented as a red-green-blue (RGB) plot, which permits visualization of the selectivity to the three intervals while varying two dimensions. As color coded in panel *A*, red represents regions of synapse space in which the Ex unit fires exclusively to the second pulse of a 50 msec IPI, but not to the 100 or 200 msec IPI; that is, a 50 msec interval detector. Similarly, green and dark blue areas represent regions of synapse space in which the Ex units respond only to the 100 or 200 ms interval, respectively. In the same manner that a computer screen makes yellow by mixing red and green, yellow in this RGB represents conditions in which the Ex unit responded to both 50 and 100 msec intervals, but not to the 200 msec interval. White areas represent regions that respond to all the intervals, but not to the first pulse. Black areas represent regions in which the cell was not interval selective: not firing at all or in response to the first pulse. The three unfilled white squares show the areas of synapse space of the traces in panel *A*. These simulations suggest that a computational function of short-term synaptic plasticity may be to allow neurons to exhibit interval selectivity and that circuits of neurons may be intrinsically capable of temporal processing. Modified from Buonomano (2000).



classes (dopaminergic antagonists and benzodiazepines) selectively interfere with second but not millisecond processing. To this author's knowledge, there have been no reports of drugs that interfere selectively with millisecond processing. Future experiments will be necessary to determine whether the above results are due to direct action on a timing mechanism or more nonspecific actions on arousal and/or cognition.

Neural Mechanisms Underlying Sensory Timing

The studies above addressed the psychophysical characteristics and localization of temporal processing, but not the actual underlying mechanisms. The term *mechanisms* refers to the neural properties that are actually sensitive to time, rather than involved in the readout. For example, looking at the readout of a watch does not necessarily provide us with any information about whether timing is occurring as a result of counting the revolution of mechanical gears or as a result of counting the oscillations of a quartz crystal. There have been a number of models of the possible neuronal mechanisms underlying timing. Rather than fully review these models, a summary of the general types of models will be provided. For simplicity, the models will be divided into two classes: labeled lines and population models. A third class is the clock model, of which internal clocks are the prototype. These models, which were described above, will not be discussed, because they are unlikely to be involved in millisecond timing, and few neurally realistic models have been put forth for them.

Labeled Lines

The majority of models that have addressed the neural mechanisms underlying timing have been influenced by the delay line model used for microsecond processing. In these models, there is an array of neurons, each of which responds selectively to a specific interval. This is considered a labeled line because there is a separate channel or neuron for each interval.

To implement labeled lines in the range of tens to hundreds of milliseconds, some temporal property must be present that allows neurons to respond selectively to a given interval. Because there must be a range of interval-selective units, whatever the time-dependent property is, there must be a spectrum of different time constants for different units. The time-dependent property can take various forms, including 1) oscillators (Fujita 1982; Miall 1989), 2) slow biochemical reactions such as the metabotropic glutamate receptor (Fiala and others 1996) or slow IPSPs combined with rebound excitation (Sullivan 1982; Margoliash 1983; Jaffe 1992), 3) intrinsic currents resulting in delayed spiking (Beggs and others 2000), and 4) cell thresholds combined with a constant rate of synaptic integration (Antón and others 1991).

What these models have in common is that in each case there are elements that are specialized for a given interval. Different elements are explicitly tuned to different intervals by adjusting the time constants, and different elements are set to different values. Because timing at different intervals is performed by independent groups of neurons, one prediction is that it is possible to abolish timing for a 250 msec interval, whereas 50 msec timing remains normal. Computationally, these models are very effective for simple tasks such as interval discrimination. However, in their simplest implementation, they are not well suited for complex forms of temporal processing such as sequences and speech.

Population Clocks

In population clocks (or population models), time is coded in the population activity of a network of neurons—any given neuron will contain little temporal information. An additional difference from labeled line models is that there is not an explicit range of time constants or time delays specifically set to capture specific intervals. Population models are a distributed type of timing; it should not be possible to create localized lesions that selectively impair one interval but not others. These models generally rely on local network dynamics and time-dependent changes in network state. The time-dependent changes in the state of the network can be the result of time-dependent properties such as short-term synaptic plasticity (Buonomano and Mauk 1994; Buonomano and Merzenich 1995), or they can be due to inhibitory feedback in local circuits (Buonomano and Mauk 1994; Mauk and Donegan 1997; Medina and others 2000).

One population model for sensory processing relies on the interaction between network dynamics and time-dependent synaptic properties (Buonomano and Merzenich 1995; Buonomano 2000)—short-term synaptic plasticity and slow synaptic events. Any initial event that arrives in a network of neurons can activate a population of neurons and will trigger a series of time-dependent properties. Thus, at the arrival of a second event 100 msec later, the same stimulus will arrive in a different network state. Due to synaptic facilitation/depression, the same synapses used 100 msec before are now stronger or weaker. Additionally, excitatory and inhibitory neurons may still be hyperpolarized by slow IPSPs. As a result, the same input can activate different populations of neurons dependent on the recent stimulus history of the network. In this type of model, a spectrum of different time constants is not present, but nevertheless neurons can respond selectively to a range of different intervals. Indeed, even in a simple network composed of two neurons it can be shown that neurons can be tuned to different intervals by changing synaptic strengths (see Box 1). Artificial network implementations of this model have been shown to be able to discriminate intervals and simple temporal sequences (Buonomano and Merzenich 1995; Buonomano 2000).

A different type of population model has been proposed to show how the cerebellar cortex may account for the timing of eye-blink conditioning (Buonomano and

Mauk 1994; Mauk and Donegan 1997; Medina and others 2000). Specifically, in the presence of a conditioned stimulus, the population activity of active granule cells changes dynamically owing to negative feedback through the granule→Golgi→granule loop. In this model, time is encoded in the population of active granule cells, and it can be read out by changing the weights of the granule-Purkinje synapses.

Conclusions

A half-century after Lashley wrote his article "The Problem of Serial Order in Behavior," the field of temporal processing is still in its infancy. However, the studies to date have allowed insights into the nature of timing. Multiple lines of evidence indicate that distinct neural mechanisms underlie millisecond and second timing. Both psychophysical and pharmacological data indicate that interval discrimination of 100 and 1000 msec tasks relies on different mechanisms, although it is not clear exactly where the boundary lies or how much overlap there is. Within the millisecond range, there is evidence that timing can undergo perceptual learning. Importantly, learning seems to generalize across modalities but not intervals. This suggests that there are central timing mechanisms in place (which does not exclude distributed timing) that are tuned to specific intervals. It is with regard to the neural mechanisms that underlie timing that relatively little progress has been made. How do neurons time external and internal events? It seems likely that the answer to this question will require an understanding of the temporal dynamics of networks of neurons. Progress is being made in recording from large numbers of neurons and analyzing the spatio-temporal patterns of activity within networks. Thus, as more neuroscientists start looking at responses to complex stimuli, and temporal discrimination tasks, we will be at last in position to make significant headway to the problem posed by Lashley 50 years ago.

References

Ackermann GS, Hertich I, Daum I. 1997. Categorical speech perception in cerebellar disorders. Brain Lang 60:323–31.

Ackermann GS, Hertich I, Daum I. 1999. Cerebellar contributions to the perception of temporal cues within the speech and nonspeech domain. Brain Lang 67:228–41.

Anderson JA, Rosenfeld E. 1988. Neurocomputing: foundations of research. Cambridge (MA): MIT Press.

Antón PS, Lynch G, Granger R. 1991. Computation of frequency-to-spatial transform by olfactory bulb glomeruli. Biol Cybern 65:407–14.

Artieda J, Pastor MA, Lacuz F, Obeso JA. 1992. Temporal discrimination is abnormal in Parkinson's disease. Brain 115:199–210.

Beggs JM, Moyer JR, McGann JP, Brown TH. 2000. Prolonged synaptic integration in perirhinal cortical neurons. J Neurophysiol 83:3294–8.

Braitenberg V. 1967. Is the cerebellar cortex a biological clock in the millisecond range? Prog Brain Res 25:334–6.

Buonomano DV. 2000. Decoding temporal information: a model based on short-term synaptic plasticity. J Neurosci 20:1129–41.

Buonomano DV, Mauk MD. 1994. Neural network model of the cerebellum: temporal discrimination and the timing of motor responses. Neural Comput 6:38–55.

Buonomano DV, Merzenich MM. 1995. Temporal information transformed into a spatial code by a neural network with realistic properties. Science 267:1028–30.

Buonomano DV, Merzenich MM. 1998. Cortical plasticity: from synapses to maps. Annu Rev Neurosci 21:149–86.

Carr CE. 1993. Processing of temporal information in the brain. Annu Rev Neurosci 16:223–43.

Creelman CD. 1962. Human discrimination of auditory duration. J Acoust Soc Am 34:582–93.

Divenyi P, Danner WF. 1977. Discrimination of time intervals marked by brief acoustic pulses of various intensities and spectra. Percept Psychophys 21:125–42.

Fiala JC, Grossberg S, Bullock D. 1996. Metabotropic glutamate receptor activation in cerebellar Purkinje cells as substrate for adaptive timing of the classically conditioned eye-blink response. J Neurosci 16:3760–34.

Fujita M. 1982. An adaptive filter model of the cerebellum. Biol Cybern 45:195–206.

Gibbon J, Malapani C, Dale CL, Gallistel CR. 1997. Toward a neurobiology of temporal cognition: advances and challenges. Curr Opin Neurobiol 7:170–84.

Grondin S, Rousseau R. 1991. Judging the duration of multimodal short empty time intervals. Percept Psychophys 49:245–56.

Harrington DL, Haaland KY, Hermanowicz N. 1998. Temporal processing in the basal ganglia. Neuropsychology 12:3–12.

Harrington DL, Haaland KY, Knight RT. 1998. Cortical networks underlying mechanisms of time perception. J Neurosci 18:1085–95.

Ivry R. 1996. The representation of temporal information in perception and motor control. Curr Opin Neurobiol 6:851–7.

Ivry RB, Keele SW. 1989. Timing functions of the cerebellum. J Cognit Neurosci 1:136–52.

Jaffe S. 1992. A neuronal model for variable response latency. In: Eeckman FH, editor. Analysis and modeling of neural systems. Boston: Kluwer. p 405–10.

Jueptner M, Rijntjes C, Weiller C, Faiss JH, Timmann D, Mueller SP, and others. 1995. Localization of a cerebellar timing process using PET. Neurology 45:1540–5.

Keele SW, Pokorny RA, Corcos DM, Ivry R. 1985. Do perception and motor production share common timing mechanisms: a correlational analysis. Acta Psychologica 60:173–91.

King DP, Takahashi JS. 2000. Molecular genetics of circadian rhythms in mammals. Annu Rev Neurosci 23:713–42.

Lashley K. 1951 [1960]. The problem of serial order in behavior. In: Beach FA, Hebb DO, Morgan CT, Nissen HW, editors. The neuropsychology of Lashley. New York: McGraw-Hill.

Lehiste I, Olive JP, Streeter LA. 1976. Role of duration in disambiguating syntactically ambiguous sentences. J Acoust Soc Am 60:1199–202.

Margoliash D. 1983. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. J Neurosci 3:133–43.

Matell MS, Meck WH. 2000. Neuropsychological mechanisms of interval timing behavior. Bioessays 22:94–103.

Mauk MD, Donegan NH. 1997. A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. Learn Mem 3:130–58.

Meck WH. 1996 Neuropharmacology of timing and time perception. Cognit Brain Res 3:227–42.

Medina JF, Garcia KS, Nores WL, Taylor NM, Mauk MD. 2000. Timing mechanisms in the cerebellum: testing predictions of a large-scale computer simulation. J Neurosci 20:5516–25.

Meegan DV, Aslin RN, Jacobs RA. 2000. Motor timing learned without motor training. Nature Neurosci 3:860–2.

Miall C. 1989. The storage of time intervals using oscillating neurons. Neural Comput 1:359–71.

Michon JA. 1985. The complete time experiencer. In: Michon JA, Jackson JL, editors. Time, mind and behavior. Berlin: Springer-Verlag. p 21–52.

Nagarajan SS, Blake DT, Wright BA, Byl N, Merzenich MM. 1998. Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. J Neurosci 18:1559–70.

Perrett SP, Ruiz BP, Mauk MD. 1993. Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses. J Neurosci 13:1708–18.
Rammsayer T. 1992. Effects of benzodiazepine-induced sedation on temporal processing. Hum Psychopharmacol 7:311–18.
Rammsayer TH. 1994. Effects of practice and signal energy on duration discrimination of brief auditory intervals. Percept Psychophys 55:454–64.
Rammsayer TH. 1997. Are there dissociable roles of the mesostriatal and mesolimbocortical dopamine systems on temporal information processing in humans? Biol Psychol/Pharmacopsychol 35:36–46.
Rammsayer TH. 1999. Neuropharmacological evidence for different timing mechanisms in humans. Q J Exp Psychol 52B:273–86.
Rammsayer TH, Lima SD. 1991. Duration discrimination of filled and empty auditory intervals: cognitive and perceptual factors. Percept Psychophys 50:565–74.
Rosseau R, Poirier J, Lemyre L. 1983. Duration discrimination of empty time intervals marked by intermodal pulses. Percept Psychophys 34:541–8.
Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. 1995. Speech recognition with primarily temporal cues. Science 270:303–4.
Spencer RM, Zelaznik HN, Ivry R. 2000. How do we time movements. Soc Neurosci Abstr 26:178.
Sullivan WE. 1982. Possible neural mechanisms of target distance coding in auditory system of the echolocating bat *Myotis lucifugis*. J Neurophysiol 48:1033–47.
Tallal P. 1994. In the perception of speech time is of the essence. In: Buzsaki G, Llinas R, Singer W, Berthoz A, Christen Y, editors. Temporal coding in the brain. Berlin: Springer-Verlag. p 291–9.
Treisman M. 1963. Temporal discrimination and the indifference interval: implications for a model of the 'internal clock'. Psychol Monogr 77:1–31.
Westheimer G. 1999. Discrimination of short time intervals by the human observer. Exp Brain Res 129:121–6.
Wright BA, Buonomano DV, Mahncke HW, Merzenich MM. 1997. Learning and generalization of auditory temporal-interval discrimination in humans. J Neurosci 17:3956–63.