In: Journal of Speech and Hearing Research,
22, pp. 311-320.

EFFECTS OF THREE SYLLABLE DURATIONS USING THE MELODIC INTONATION THERAPY TECHNIQUE

SUSAN A. LAUGHLIN *and* MARGARET A. NAESER

Veterans Administration Hospital, Palo Alto, California

WILLIAM P. CORDON

Foothill College, Los Altos Hills, California

Five nonfluent aphasics were presented English phrases with three syllable durations; a regular speech, non-intoned duration < 1 sec per syllable, and two modified Melodic Intonation Therapy (MIT) intoned durations of 1.5 sec per syllable and 2.0 sec per syllable. All subjects had the greatest number of correct phrase productions at the longest MIT duration ($p < 0.001$) and the greatest number of failures at the regular non-intoned duration. Therefore, syllable duration is an important acoustic parameter to consider when using an MIT technique with severe nonfluent aphasics.

Moderate to severe limitation in speech production is a prominent characteristic of the nonfluent aphasias. The four nonfluent aphasia syndromes recognized by some aphasiologists include Transcortical Motor aphasia, Broca's aphasia, Mixed aphasia, and Global aphasia. Transcortical Motor aphasia presents relatively intact repetition performance when compared to propositional speech, but in the Broca's, Mixed, and Global aphasia syndromes conversational speech and phrase repetition are similarly nonfluent. Although the nonfluent aphasic patient's speech has halting and effortful initiation, awkward articulation, reduced melodic line, and telegraphic grammatical constructions, many of these patients reportedly can sing familiar songs with good articulation (Geschwind, 1971). Some treatment methods for nonfluent aphasia have emphasized the singing of propositional speech in a clinical setting (Gerstman, 1964; Keith and Aronson, 1975). The use of Melodic Intonation Therapy (MIT) as a successful treatment technique to enhance formulation of propositional language with nonfluent aphasics was first introduced by Albert, Sparks, and Helm (1973), and further examined by Sparks, Helm, and Albert (1974).

Sparks and Holland (1976) state that the MIT technique utilizes "intoned patterns" which contain three elements—melodic line, rhythm, and points of stress. The linguistic concept "stress" is described acoustically in terms of intensity, spectrum, fundamental frequency, and duration (Malmberg, 1968). Fry (1955) found that with normal subjects increased duration is the single

most important factor in perception of stress in English syllables. Other investigators have concluded that for aphasic patients with left hemisphere lesions and auditory processing defects, expanded duration of both verbal and nonverbal acoustic stimuli improves performance on a variety of experimental tasks requiring temporal resolution and sequencing analysis (Efron, 1963; Albert and Bear, 1973; Tallal and Newcombe, 1978).

In this study, we manipulated syllable duration and measured its effect on phrase production ability in a small sample of nonfluent aphasic subjects. Subjects produced words and phrases at three syllable durations, a regular speech duration (with an approximate mean of 0.28 sec) and two expanded MIT intoned durations (1.5 and 2.0 sec). The purpose of this study was to determine if a prolonged syllable duration would improve subjects' verbal performance within the MIT technique.

METHOD

Subjects

Five right-handed male CVA patients were selected as subjects for this study. Four of these subjects had left hemisphere vascular lesions and one subject had bilateral lesions. These lesions were verified from localizing neurologic signs and CT scans. CT scans were done 7 to 108 months post-onset, and were interpreted by a staff radiologist. The composites in Figure 1 show that all five subjects had lesions in the left hemisphere. One of the five had an additional small lesion in the right occipital lobe, but like the others, his predominant lesion was in the distribution of the left middle cerebral artery. Subjects ranged in age from 47 to 67 years, with a mean of 58 years. Months post-onset ranged from 4 to 108, with a mean of 29 months.

Subject selection also was based on the following clinical criteria:

(1) Patients were classified as nonfluent in conversational speech on the *Boston Diagnostic Aphasia Examination* Rating Scale Profile of Speech Characteristics (Goodglass and Kaplan, 1972). Additional language testing placed each patient within a general aphasia syndrome. Two of the five patients were classified as presenting Global aphasia, two were classified as presenting Mixed aphasia, and one was classified as presenting Broca's aphasia. None of the five patients in the sample exhibited Transcortical Motor aphasia, and all five had moderate to severe impairments on high probability phrase repetition. Descriptive information on the subjects is provided in Table 1. See Figure 2 for composite range of speech profile ratings.

(2) A patient's pre-experimental therapy experience with MIT was limited to less than one month.

Stimulus Materials

Sixty stimulus phrases and single word items were administered by live

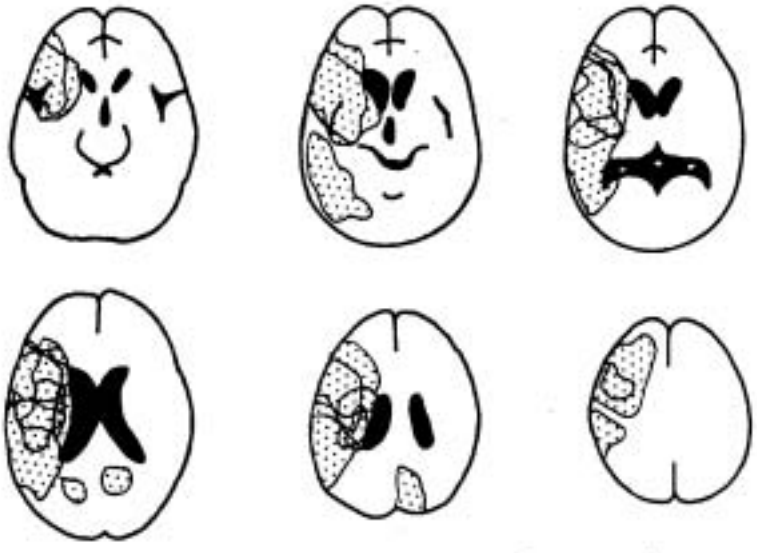

FIGURE 1. Six CT scan slices showing superimposed composite lesion sites for five subjects with nonfluent aphasia. Moving from left to right, top to bottom row, schematics represent CT slices of ascending level within the brain. Stipled areas mark presence of infarcted brain tissue. At upper left slice, Broca's area may be visualized on periphery just lateral to anterior horn of left lateral ventricle (two subjects had identified lesions on this slice), and moving across the top row of CT slices (and progressing superiorly through the brain), the middle slice represents both Broca's and Wernicke's language areas; Wernicke's area is represented by the top right slice. Bottom row on left is level of supramarginal gyrus (superior to Wernicke's slice in the brain), where all 5 subjects had identified lesions. Bottom right is uppermost CT slice of those shown.

TABLE 1. Descriptive information on age, months post-onset, auditory comprehension, and aphasia type for 5 subjects.

S	Age	Months Post-Onset	*Auditory Comprehension Z-Score (Boston Diagnostic Aphasia Examination)*	*% Correct Token Test Score (Spreen & Benton) 1969*	*Aphasia Type*
R.B.	60	6	−1.95	48	Global
J.A.	55	4	−2.10	48	Global
D.T.	47	24	−0.20	50	Mixed
J.B.	67	4	+0.89	88	Broca's
C.B.	64	108	+0.46	75	Mixed

voice to each subject, 20 phrases per duration condition. The experimenter (SL) produced phrases with live voice because the MIT technique does not use pre-recorded stimulus materials. In addition, aphasic individuals using this technique sometimes require repetitions and back-up procedures following error productions, necessitating some flexibility in stimulus presentation. These stimulus phrases and word items consisted of spondee nouns (baseball), social greetings (how are you), actions (smoking), or verb-object commands (open

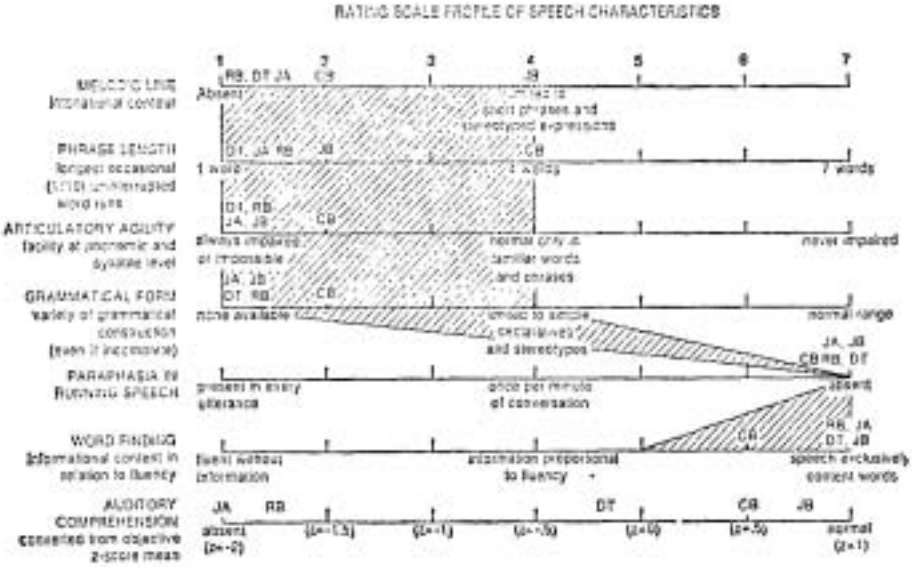

FIGURE 2. Composite range of speech profile ratings characteristic of five nonfluent aphasic subjects. Profile outline marks the typical range of scores for a nonfluent aphasia, excluding auditory comprehension, which may range from a mild to a severe deficit.

the door). With the 1.5 and 2.0 sec intoned durations, stressed syllables received high tone, unstressed syllables received low tone. The frequency difference between high and low tones approximated the musical interval of a fourth. Stimulus phrases were ordered to avoid consecutive duplication of tone contours, to reduce the tendency for subjects to perseverate tone contours and syllables from one phrase to the next. The stimulus phrases and words were not balanced for phonologic or syntactic complexity.

To establish an approximate value for mean vowel duration in her regular, non-intoned speech, the experimenter recorded 16 CVC syllables, representing the 12 vowels and four diphthongs in English. Each syllable was produced with primary stress. Sound spectrograms were made of these syllables (Kay Elemetrics Sonagraph Model 6061-B) and the vowel durations were measured (Naeser, 1970). The durations observed ranged from 128 msec to 454 msec, and the mean was 282 msec. This mean vowel duration for regular speech, non-intoned syllables was obtained as an approximate baseline for subsequent comparison with the prolonged intoned MIT durations.

The 1.5 sec and 2.0 sec intoned syllable durations were timed, while they were being produced by the experimenter, with an optical metronome. The metronome was calibrated to flash a 2-watt light once every 1.5 or 2.0 sec (it was visible only to the experimenter during the experimental sessions).

Three sound spectrograms for the sample phrase, *good-bye*, as produced by the experimenter at the three different syllable durations, are shown in Figure 3. The measured vowel durations (msec) were the following:

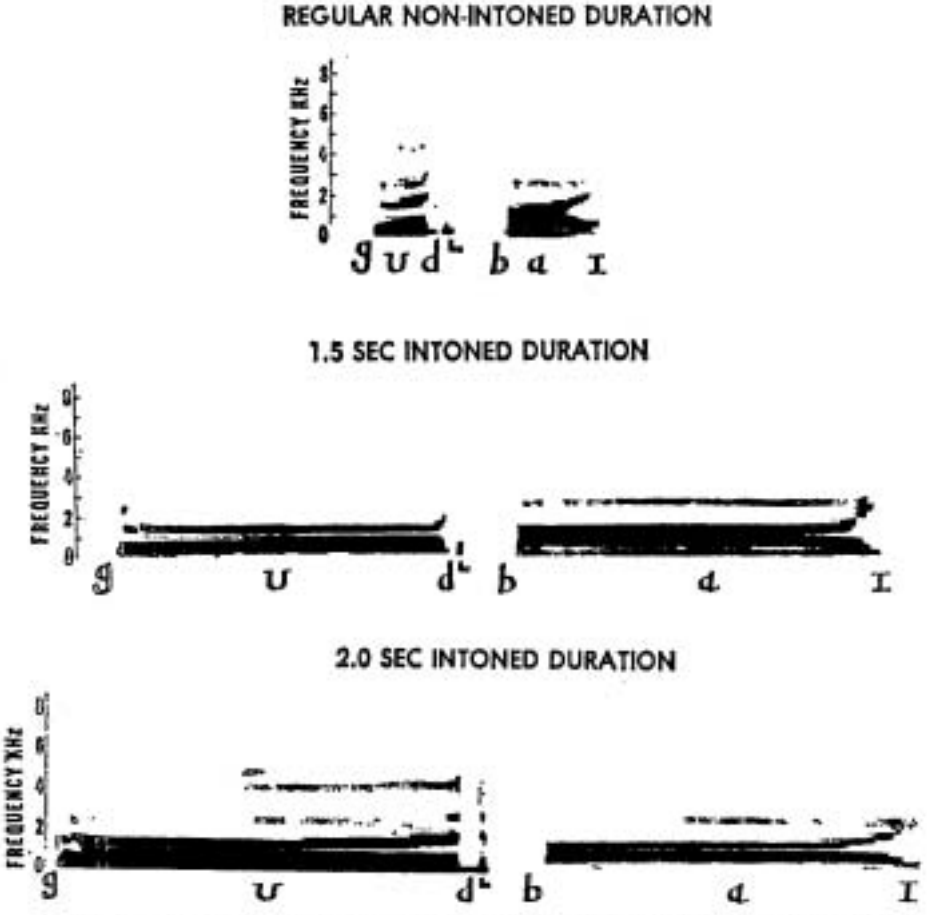

FIGURE 3. Sound spectrograms for the phrase, "good-bye", produced live voice at three syllable durations: regular speech, non-intoned duration (285 msec /ʊ/, 577 msec /aɪ/), 1.5 sec intoned duration (1222 msec /ʊ/, 1642 msec /aɪ/), 2.0 sec intoned duration (1860 msec /ʊ/, 1935 msec /aɪ/).

(1) regular non-intoned duration: 285 /ʊ/, 577 /aɪ/
(2) 1.5 sec intoned duration: 1222 /ʊ/, 1642 /aɪ/
(3) 2.0 sec intoned duration: 1860 /ʊ/, 1935 /aɪ/

Thus, all durations were approximate. In no instance did the regular speech, non-intoned duration equal the 1.5 or 2.0 sec intoned MIT syllable durations, and the 1.5 and 2.0 sec intoned syllable durations did not overlap.

Procedure

A baseline phrase repetition level (number of syllables per phrase which could be reliably repeated) was established for each subject. It should be noted that a baseline for number of syllables per phrase that a patient can repeat is not a standard MIT pre-therapy item. The phrase repetition baselines

were determined using a laddering procedure, that is, a correct phrase repetition meant that the next phrase given was one syllable longer; incorrect phrase repetition meant that the next phrase given was one syllable shorter. The baseline test phrases used were representative of, but not identical to, those used in experimental sessions. Thus, each individual subject was assigned appropriate phrase lengths adjusted to the severity of his repetition impairment.

Stimulus phrase presentation in the 1.5 and 2.0 sec intoned MIT conditions followed basic procedures described in the methodology manual (Sparks, 1974)[1]. These procedures are similar to Level III, Sparks and Holland (1976). The five MIT steps are summarized as follows:

> (1) unison humming, (2) unison intoned phrase production, (3) clinician fades as patient attempts intoned production alone, (4) patient repeats intoned production alone, and (5) patient intones appropriate response to clinician's intoned question

Thus, MIT steps 1, 2, and 3 require the patient to intone phrases with the clinician, and steps 4 and 5 require the patient to intone the phrase independent of the clinician. Stimulus phrase presentation in the regular non-intoned duration condition followed this same format; however, step 1 was omitted. When non-intoned stimulus phrases were presented by the experimenter, non-intoned response phrases were expected from the subjects.

Sixty different stimulus phrases, 20 in each syllable duration condition, were administered over 12 sessions, 3 sessions per week. In one week, a subject was given three 30-minute sessions. In a single session, five phrases were presented in only one of three syllable duration conditions. A given phrase was never used with more than one syllable duration condition. Each duration condition (regular non-intoned, 1.5 sec intoned, or 2.0 sec intoned) was presented in one session per week, and the order of weekly presentation was counterbalanced.

RESULTS

Table 2 presents results for each subject, including length of stimulus phrase

TABLE 2. Response scores for three syllable durations at MIT steps 1-5.

		*Regular Non-Intoned Duration MIT Steps**				*1.5 Sec Intoned Duration MIT Steps**					*2.0 Sec Intoned Duration MIT Steps**				
S	*Length Of Stimulus Phrase In Syllables*	*2*	*3*	*4*	*5*	*1*	*2*	*3*	*4*	*5*	*1*	*2*	*3*	*4*	*5*
R.B.	2	4	4	4	3	20	12	11	9	8	20	12	12	11	10
J.A.	2	8	8	6	6	20	13	11	11	10	20	17	17	17	17
D.T.	2-3	1	1	1	1	20	20	19	14	13	20	19	16	16	16
J.B.	4-5	3	3	3	3	20	20	20	17	14	20	20	20	19	19
C.B.	5-6	6	6	6	6	20	20	20	13	12	20	20	20	20	20
	Mean	4.7	4.4	4.0	3.8	20	17	16.2	12.8	11.4	20	17.6	17	16.6	16.4

*Maximum score at each step is 20.

[1]Sparks, Melodic Intonation Therapy: Methodology Manual, Unpublished manual (1979).

and the number of phrases (out of 20) each subject produced correctly at every MIT step, within each syllable duration condition. Note that MIT step 1 (patient hums melody) was not used with the regular non-intoned duration condition. Mean correct phrase production across steps 2–5, in the regular non-intoned condition was 4.15 (20.75%); in the 1.5 sec per syllable intoned condition 14.35 (71.75%); and in the 2.0 sec per syllable intoned condition 16.90 (84.5%). Thus, all subjects had more correct phrase productions with the prolonged intoned conditions, than with the regular non-intoned condition.

As the difficulty of MIT steps increased, correct phrase production decreased across all syllable duration conditions. The percent correct phrase production for each syllable duration condition, at each MIT step, is graphed in Figure 4. Correct phrase production with the regular non-intoned condition decreased from 22% at step 2, to 19% at step 5. With the 1.5 sec per syllable intoned condition it decreased from 83% at step 2, to 54% at step 5. With the 2.0 sec per syllable intoned condition, it decreased from 88% at step 2, to 82% at step 5. However, 2.0 sec per syllable condition was associated with the highest phrase production performance at all MIT steps.

A Repeated Measures Two-Factor Analysis of Variance was used to analyze the data further, and a summary of this analysis is presented in Table 3. The data indicate the following:

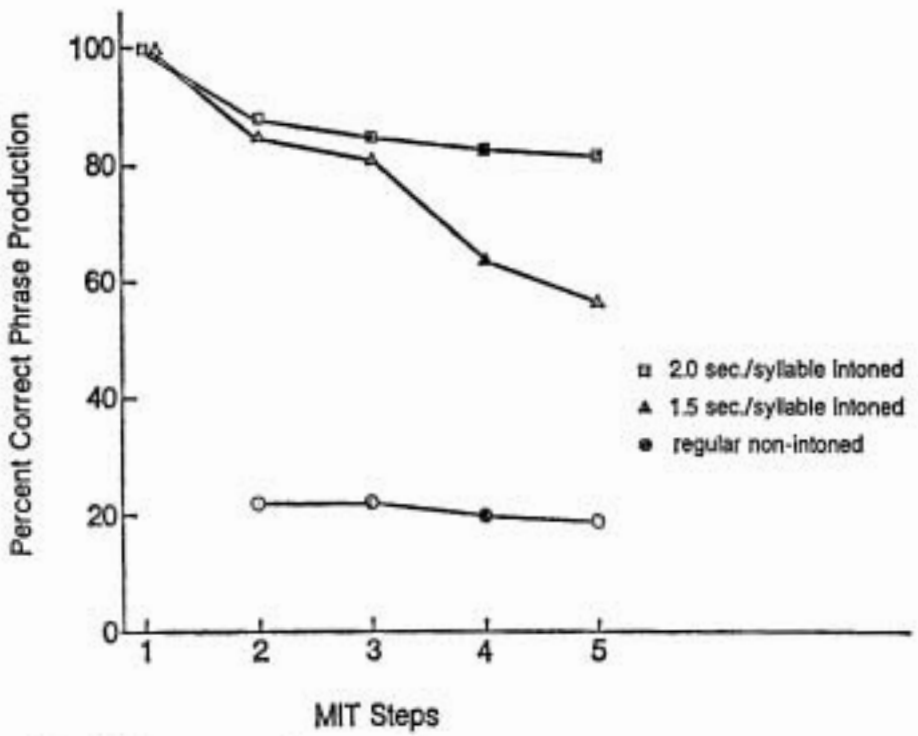

FIGURE 4. Percent correct phrase production at MIT steps 1-5, for three syllable duration conditions.

TABLE 3. Analysis of variance results (F ratios) used to evaluate differences between duration conditions, MIT steps, and interactions between duration conditions x MIT steps.

Source	*df*	*MS*	*F*	*P*
Total	59	–	–	–
Subjects	4	–	–	–
Duration Condition	2	910.35	30.55	0.001
MIT Steps	3	20.08	9.04	0.005
Duration x Steps	6	8.83	7.42	0.001
Error Durations	8	29.80	–	–
Error Steps	12	2.22	–	–
Error Durations x Steps	24	1.19	–	–

(1) phrase production was significantly affected by the syllable duration ($F = 30.55$, $p < 0.001$), (2) phrase production was significantly affected by MIT step ($F = 9.04$, $p < 0.005$), and (3) there was a significant interaction between syllable duration and MIT step ($F = 7.42$, $p < 0.001$)

Mean comparisons for duration and MIT step, and duration x MIT step interactions were carried out using Scheffe's Test (Glass and Stanley, 1970).

Overall, phrase production was significantly better at the 1.5 sec per syllable intoned condition than for the regular speech, non-intoned condition ($p < 0.005$), and for the 2.0 sec per syllable intoned condition, than at the regular non-intoned condition ($p < 0.001$). No significant difference existed between the 1.5 and 2.0 sec per syllable conditions, across all treatment conditions ($p > 0.10$). Statistically significant mean differences for each MIT step across all duration conditions were found. These mean differences were between steps 2 and 4 ($p < 0.05$), 2 and 5 ($p < 0.01$), and 3 and 5 ($p < 0.05$).

Mean comparisons were analyzed between each MIT step and the three duration conditions. All mean comparisons for the regular non-intoned condition versus the 1.5 sec per syllable intoned condition for steps 2, 3, 4, and 5 were statistically significant ($p < 0.001$). The mean comparisons between step and syllable duration were not significant for the 1.5 versus 2.0 sec per syllable durations with steps 2 and 3 ($p > 0.05$), however, the mean comparisons were significant for steps 4 ($p < 0.05$) and 5 ($p < 0.01$). Thus, phrase production with the 2.0 sec per syllable duration was significantly better than phrase production with the 1.5 sec per syllable duration only at MIT steps 4 and 5.

Summary of Results

(1) All subjects, whether moderate or severe aphasics, had significantly more correct phrase productions with the prolonged intoned conditions of 1.5 and 2.0 sec per syllable, than with the regular speech, non-intoned condition.

(2) The highest number of correct phrase productions was observed with the 2.0 sec per syllable condition, the next highest with the 1.5 sec per syllable condition, and the lowest with the regular speech, non-intoned condition.

(3) As MIT steps became more difficult, the percentage of correct phrase productions decreased, across all syllable duration conditions. Significantly

more correct phrase productions existed at steps 2 (unison phrase intonation) and 3 (clinician fades participation), than at steps 4 (solo phrase production) and 5 (solo production in response to the clinician's related question).

(4) The interaction between syllable duration and MIT step indicated that phrase production performance with the 2.0 sec/syllable duration was significantly better than the 1.5 sec/syllable duration only at steps 4 and 5.

DISCUSSION

This is the first attempt to investigate the effect of increased syllable duration on Melodic Intonation Therapy. The findings show that significantly more accurate phrase production performance existed at all steps for prolonged MIT durations. Further, performance at MIT steps 4 and 5 (when the subject intoned the phrase without verbal assistance from the experimenter) was significantly more accurate with the 2.0 than with the 1.5 sec per syllable duration condition. The nonfluent subjects who participated in this study benefitted not only from the 1.5 sec per syllable prolonged intoned duration, but also from the additional 500 msec at the 2.0 sec per syllable prolonged intoned duration. Our findings agree with previously cited research regarding the facilitating effect of expanded speech duration. More specifically, recommended procedural adaptations may be in order concerning syllable duration within the MIT technique.

The goal of the MIT technique is to facilitate formulation of propositional language. This goal may be attained with certain nonfluent aphasics after a given period of therapy with MIT at Levels 1, 2, 3, and 4 (Sparks and Holland, 1976). A criterion of 90 per cent correct phrase production performance on 10 consecutive scores has been set for a patient to progress from one level to the next. Results from this study would suggest that if patients are unable to reach criterion at step 5 with shorter MIT durations (< 2.0 sec per syllable), the patient would benefit from the prolonged intoned duration.

In this study, every syllable (stressed and unstressed) in the MIT intoned conditions was prolonged to 1.5 or 2.0 sec. This does not conform to the stress patterns of normal English. However, it was used in this study to ascertain whether or not increased duration facilitated improved verbal performance. The 2.0 sec per syllable duration did facilitate significantly more correct responses at MIT steps 4 and 5 (Level 3). Further study is recommended to determine whether the same effect can be obtained when increased syllable duration of 2.0 sec is used only on the normally stressed syllables.

ACKNOWLEDGMENT

This research was completed while the senior author was at San Jose State University and is based on a master's thesis submitted to the graduate faculty of that institution. The authors wish to thank the following for their valuable contribution: Katharine G. Butler, Dean of Graduate Studies, San Jose State University; Mary V. Dickerson, Associate Professor, Special Education Department, San Jose State University; and Ted Montemurro,

Associate Professor, Special Education Department, San Jose State University. The authors also wish to acknowledge Leslie M. Zatz, M.D. and Robert W. Hayward, M.D., from the Radiology Service, Palo Alto V.A. Hospital, for their assistance with the CT scan lesion analysis. Requests for reprints should be sent to Susan Laughlin, Audiology/Speech Pathology Service, Veterans Administration Hospital (126), Palo Alto, California 94304. Margaret A. Naeser is presently with the Aphasia Research Unit, Boston Veterans Administration Hospital, Boston, Massachusetts.

REFERENCES

ALBERT, M., SPARKS, R., and HELM, N., Melodic Intonation Therapy for aphasia. *Arch. Neurol.*, **29**, 130-131 (1973).

ALBERT, M., and BEAR, D., Time to understand: a case study of word deafness with reference to the role of time in auditory comprehension. *Brain*, **97**, 373-384 (1973).

EFRON, R., Temporal perception, aphasia and Deja Vu. *Brain*, **86**, 403-424 (1963).

FRY, D. B., Duration and intensity as physical correlates of linguistic stress. *J. acoust. Soc. Am.*, **27**, 765-768 (1955).

GERSTMAN, H., A case of aphasia. *J. Speech Hearing Dis.*, **29**, 89-91 (1964).

GESCHWIND, N., Current concepts: aphasia. *New Eng. J. Med.*, **284**, 654-656 (1971).

GLASS, G., and STANLEY, J., *Statistical Methods in Education and Psychology*. Englewood Cliffs, New Jersey: Prentice-Hall (1970).

GOODGLASS, H., and KAPLAN, E., *The Assessment of Aphasia and Related Disorders*. Philadelphia: Lea & Febiger (1972).

KEITH, R., and ARONSON, A., Singing as therapy for apraxia of speech and aphasia: report of a case. *Brain Lang.*, **2**, 483-488 (1975).

MALMBERG, B. (ed.), *Manual of Phonetics*. New York: North-Holland (1968).

NAESER, M., *Criteria for Segmentation of Vowels on Duplex Oscillograms*. Madison, Wisconsin: Wisconsin Research and Development Center for Cognitive Learning, Technical Report #124 (1970).

SPARKS, R., HELM, N., and ALBERT, M., Aphasia rehabilitation resulting from Melodic Intonation Therapy. *Cortex*, **10**, 303-316 (1974).

SPARKS, R., and HOLLAND, A., Method: Melodic Intonation Therapy for aphasia. *J. Speech Hearing Dis.*, **41**, 287-297 (1976).

SPREEN, O., and BENTON, A. L., *Token Test Instructions Long Form*. University of Victoria, Victoria, B.C.: Neuropsychology Laboratory, Department of Psychology (1969).

TALLAL, P., and NEWCOMBE, F., Impairment of auditory perception and language comprehension in dysphasia. *Brain Lang.*, **5**, 13-24 (1978).

Received April 17, 1978.
Accepted October 17, 1978.