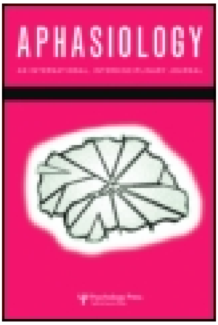

Aphasiology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/paph20

Effects of rate and rhythm control treatment on consonant production accuracy in apraxia of speech

Julie L. Wambaugh [a] & Aida L. Martinez [b]
[a] University of Utah, Salt Lake City, USA and VAMC, Salt Lake City, USA
[b] VAMC, Salt Lake City, USA
Published online: 31 Aug 2010.

APHASIOLOGY, 2000, VOL. 14, NO. 8, 851–871

Effects of rate and rhythm control treatment on consonant production accuracy in apraxia of speech

JULIE L. WAMBAUGH[1,2]
and AIDA L. MARTINEZ[2]

[1] University of Utah, Salt Lake City, USA
[2] VAMC, Salt Lake City, USA

(*Received 29 August 1998; accepted 16 August 1999*)

Abstract

A speaker with apraxia of speech and aphasia was trained to produce multisyllabic words using a combination of metronomic rate control and hand-tapping. A multiple baseline design was used to examine the effects of treatment on sound production. Treatment was applied to three syllable words with primary stress on the first syllable while generalization was measured to: (1) untrained exemplars; (2) three syllable words with different stress patterns; (3) four syllable words; and (4) s-blend words. Positive sound changes were noted for trained and untrained words. Treatment was extended to a second set of words to which generalization had been incomplete and additional improvement was observed.

Introduction

Apraxia of speech (AOS) is a neurogenic sensorimotor speech disorder that is characterized by disturbances in articulation and prosody. This disorder is thought to occur as a result of "inefficiencies in the translation of a well-formed and filled phonologic frame to previously learned kinematic parameters assembled for carrying out the intended movement" (McNeil *et al*., 1997, p. 329). Intra and interarticulator timing has frequently been found to be compromised in AOS (Fromm *et al*. 1982, Hardcastle 1987, Itoh *et al*. 1980). This disruption in speech timing, along with spatial targeting imprecision, appears to account for the majority of aberrant speech behaviours observed in AOS.

The notion that the timing of speech production is disrupted in AOS has been reflected in a number of treatment approaches designed to control the rate or rhythm of speech production with apraxic speakers. Strategies such as metronomic pacing (Dworkin and Abkarian 1996, Dworkin *et al*. 1988), finger-tapping (Simmons 1978), prolonged speaking (Southwood 1987), vibro-tactile stimulation (Rubow *et al*. 1982), and melodic intonation therapy (MIT, Sparks and Deck 1994) have been advocated for use in the treatment of AOS and appear to have promise in improving apraxic speech production.

Shane and Darley (1978) first studied the effects of auditory rhythmic stimulation

Address correspondence to: Julie Wambaugh, Department of CMDIS, 1201 Behavioural Sciences Building, 1530 E. 390 S., University of Utah, Salt Lake City, Utah 84112, USA. Email: julie.wambaugh@health.utah.edu

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

(i.e. use of a metronome) on the speech production of apraxic speakers. In this investigation, an oral reading task was utilized to measure the effects of three imposed speaking rates on the number and types of sound errors produced. Eight apraxic subjects were asked to read paragraphs accompanied by a metronome set at their typical reading rate, a rate 25% faster than the typical rate, and a rate 25% slower than the typical rate. Additionally, they were asked to read a paragraph with no external pacing provided. The researchers found no significant differences in numbers and types of errors across the reading conditions.

In contrast to Shane and Darley's (1978) facilitation study, Dworkin and colleagues (1988, 1996) found positive effects of metronomic pacing in two treatment investigations. Dworkin *et al*. (1988) used a metronome to pace four treatment activities with a single, chronically apraxic speaker. The activities included repeated productions of the following behaviours: (1) nonspeech repetitive tongue movements; (2) alternate motion rate (AMR) syllable production; (3) multisyllabic word production; and (4) sentence production. The treatment involved treating the behaviours sequentially, in a procession from motorically least to most complex and systematically increasing the speed of production. A multiple probe design was employed and behavioural change was measured in terms of presence or absence of AOS symptoms in probes as well as in treatment sessions. Positive changes were noted in all measured behaviours following the application of treatment, with no generalization across behaviours. A stimulus generalization measure, discourse production in response to predetermined questions, revealed increased positive ratings of speech production.

In a similar investigation, Dworkin and Abkarian (1996) paired a metronome with several vowel production activities to treat a patient with chronic AOS and unilateral upper motor neuron dysarthria. The patient in this investigation demonstrated severe disruption of phonatory control along with disturbances of articulation and prosody. Treatment entailed repeated practice of isolated vowels, three-vowel sequences, and alternating vowel and /h/ sequences with metronomic pacing being gradually increased in rate throughout the course of treatment. Treatment effects were measured in terms of volitional production of nondisrupted voicing in time to the metronome's beat. Positive treatment effects were observed for all treated behaviours and some response generalization to more complex targets was noted.

Instead of providing auditory stimulation to control rate and rhythm, Rubow *et al*. (1982) employed vibrotactile stimulation with a single subject. The vibrotactile stimulation consisted of applications of a 50 Hz vibration to the subject's index finger, with the applications corresponding to each syllable of three-syllable words. An alternating treatments design was utilized to compare the effects of the experimental treatment and imitation treatment on production of multisyllabic words. Results revealed a greater improvement for words that received the vibrotactile stimulation treatment than for words that received the imitation treatment, as determined by ratings on a 16 point multidimensional scale. Unfortunately, the word lists may not have been equated in terms of difficulty of acquisition and replications of this investigation have not been conducted.

Another form of rate and rhythm control that has been advocated for use with AOS is the use of finger-counting (Simmons 1978). As employed by Simmons, this approach involved producing each word in a simple sentence while simultaneously raising a finger (as in counting) as each word was produced. The treatment also utilized written word stimuli, clinician modelling of sentence production paired with finger-counting, and unison productions with the clinician. Treatment effects in this case study with a chronic apraxic and aphasic speaker were measured by changes in verbal scores on the Porch Index of Communicative Ability (PICA, Porch 1967) and substantial positive changes were reported. A similar approach that has also been suggested for treatment of AOS patients (Square and Martin 1994, Wertz *et al*. 1984) entails pairing speech production with finger/hand movement over a pacing board (Helm 1979). Currently, there are no published empirical data regarding the use of pacing boards by AOS speakers, nor are there additional investigations of finger-counting effectiveness.

Gestural rhythmic stimulation is also part of MIT (Sparks and Deck 1994). The effects of MIT on apraxic speech production have not been directly examined. However, several nonfluent aphasic patients with impaired articulation and prosody have been observed to benefit from this treatment (Albert *et al*. 1973, Naeser and Helm-Estabrooks 1985). Although there are currently no data indicating which parameters of MIT are critical to improvements in productive speech and language, the rhythm and pacing aspects of this treatment may play a particularly important role for aphasic patients with co-occurring AOS.

The auditory, vibrotactile, and gestural rhythmic stimulation techniques discussed previously, served to regulate both the rate and rhythm of speech production. A strategy in which only the rate of speech was controlled was studied by Southwood (1987). Two apraxic speakers were instructed to use a prolonged manner of speech production while orally reading passages presented at controlled rates. The investigators found that the number of speech sound errors was reduced for both speakers as rate of speaking decreased. However, no generalization to discourse was evident and the subjects' speech productions were described as sounding unnatural.

As with most approaches to treatment for AOS, the effects of rhythm and rate control strategies appear to be promising, but lack of replications serve to restrict the degree to which such treatments may be advocated for use with AOS speakers (see Wambaugh and Doyle 1994 for a review). Additionally, it has been assumed that treatments focused upon rate and rhythm will result in improved sound productions with apraxic speakers (Square and Martin 1994). However, the effects of rhythmic stimulation on sound production accuracy has not been directly examined in a treatment paradigm. That is, investigators have not used phonetic transcription, but have used general measures of speech production (i.e. PICA scores, multidimensional ratings, presence or absence of apraxic behaviours) as indicators of treatment effects. Furthermore, in the two facilitation investigations in which sound production accuracy was measured (Shane and Darley 1978, Southwood 1987), conflicting findings were reported for the different rate control techniques of metronomic pacing and prolonged speaking.

The purpose of the present investigation was to study the effects of a rate and rhythm control treatment on sound production with an apraxic and aphasic speaker. Specifically, an apraxic speaker was trained to produce multisyllabic words utilizing a combined metronomic pacing and hand-tapping strategy. Acquisition, response generalization, and maintenance effects of treatment were examined in terms of accuracy of consonant and word production. Additionally, temporal-acoustic measures were conducted to assist in analysing potential changes in sound productions and response latencies across the course of the investigation.

Method

Subject

The participant in this investigation was a 38 year old male who was 79 months post-onset of a single, left-hemisphere hemorrhagic stroke. Radiological reports indicated a cerebral infarct in the distribution of the left middle cerebral artery. According to medical records and self-report, the participant had no prior history of speech/language disturbances, neurological problems, or psychiatric disturbances. He passed a pure-tone hearing screening at 500, 1K, 2K, and 4K Hz at 25 dB HL. The participant demonstrated mild to moderate AOS, with speech behaviours being consistent with the characteristics of AOS discussed by McNeil *et al*. (1997). Specifically, he demonstrated slow rate, an inability to increase rate and simultaneously maintain phonemic integrity, numerous sound errors with a predominance of distortions, relatively consistent errors in terms of location and type, occasional schwa intrusions, and increasing frequency of errors on words of increasing length. These behaviours were observed in conversational, narrative, and procedural discourse as well as in repetition and structured speech elicitation tasks. Examples of the speaker's misarticulated productions of experimental stimuli (i.e. multisyllabic words) are presented in appendix 1.

The speaker exhibited no significant abnormalities of muscle tone or strength or any dysarthria as discussed by Duffy (1995). He also evidenced Broca's aphasia, as classified by the Western Aphasia Battery (Kertesz 1982). His productive spoken language was characteristic of agrammatic aphasia (Saffran *et al*. 1989) and was typified by a predominance of nouns, a lack of functors, and relatively infrequent verb usage. Assessment results obtained prior to treatment are presented in table 1.

Experimental stimuli

One hundred and twenty words were used to measure the effects of treatment on sound production. These words comprised five groups: (1) group 1—40 three syllable words with primary stress on the first syllable (e.g. wonderful); (2) group 2—40 three syllable words with primary stress on the second syllable (e.g. inherit); (3) group 3A—10 three syllable words with primary stress on the third syllable (e.g. guarantee); (4) group 3B—20 four-syllable words (e.g. television); and (5) group 3C—10 words beginning with clusters of 3 consonants (e.g. spring) (see appendix 2). One-half of the words in groups 1 and 3A were designated as treatment items and the remaining half in each group were designated as response generalization items. The assignment of items to treatment or no-treatment groups was quasi-random in that number of total consonants across groups was balanced, with selection being otherwise randomized. The items in groups 2, 3B, and 3C were not divided into treated and untreated groups because these groups never received treatment.

Experimental design

A multiple baseline design across behaviours was used to examine treatment effects. The accuracy of consonant production in the experimental words served as the dependent variable. Baseline probes were conducted until correct responding was

Table 1. Results of speech and language assessment battery

Measure	Result
Western Aphasia Battery (Kertesz 1982)	
(aphasia quotient—100 possible)	74.4
(aphasia classification)	Broca's
Test of Adolescent/Adult Word-Finding (German 1990)	
(raw score from brief test—40 possible)	31
Apraxia Battery for Adults (Dabul 1979)	mild &
(majority of ratings)	moderate
Discourse Measures	
MLU	1.29
# different word roots	20
% CIUs	65
Assessment of Intelligibility of Dysarthric Speech—Words (Yorkston and Beukelman 1981)	
(% intelligible)	80%

either stable or descending across all groups of words. Treatment was then applied first to group 1 words (three syllable, stress on first syllable) while probing continued with all groups of words. Treatment was then applied to group 3A words (three syllable, stress on the third syllable) while probing continued.

Baseline

During baseline, correct production of sounds in the 120 experimental words was measured in word repetition probes. The 120 words were presented in three groups of 40 words each (i.e. group 1, group 2, and groups 3A, 3B, 3C combined, as described previously), with the order of groups and the order of words within groups randomized. The experimenter produced each word and asked the speaker to repeat the word as accurately as possible. No feedback was provided regarding accuracy of production. Five baseline probes were conducted.

Treatment

During the treatment phase, the speaker was trained to produce three syllable words in rhythm with a metronome and in conjunction with hand-tapping. Treatment was applied first to the twenty treatment words in group 1 (three syllable, primary stress on first syllable). Treatment involved the use of a metronome (audible click plus small flashing red light) set to 93 beats per minute (bpm). A three syllable word produced at rate of 93 beats per minute resulted in a word duration of about 2 seconds. This word duration was approximately 50% longer than the average three syllable word duration produced by the speaker during baseline (i.e. approximately 1300 ms per word). The specific value of 93 bpm was determined by allowing the participant to select the rate from a range of 90 to 100 bpm.

The speaker was trained to produce one syllable per beat while tapping his hand in unison with the metronome. Clinician participation during production was systematically faded. Additionally, rate was gradually increased to approximate a more normal rate. 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.

During each treatment session, treatment at the current level of the treatment hierarchy was applied to every treatment word (e.g. if criterion for level 1 had not yet been reached, then all treatment words were submitted to the treatment steps in level 1). Each of the treatment words was entered into treatment in random order, until the entire group had been presented. This process was then repeated as frequently as possible within the treatment session (see appendix 3 for entire treatment approach).

Probes, identical to those conducted in baseline, were administered preceding every two to three treatment sessions. Treatment was conducted three times per week in an outpatient clinic by the authors. Administration of treatment was alternated between the authors, with the first author conducting approximately every third to fourth session. Sessions were approximately one hour in length including probes.

Treatment continued with group 1 words until at least 90% accuracy of consonant production (i.e. not word production) was evident in trained words in three consecutive probe sessions. Treatment was then extended to group 3A words (three syllable, stress on third syllable). Treatment was not extended to group 2 words because correct consonant production had already reached levels approaching 90%.

Maintenance and follow-up

Maintenance of previously trained behaviours (i.e. group 1) was measured during training of group 3 items. A follow-up probe, in which production of items in all groups was examined, was conducted at 3 weeks following completion of all training.

Scoring

All productions were transcribed on-line, using broad phonetic transcription. All baseline and probe sessions were recorded and these audiorecordings were used to verify transcriptions. Each group of words in every probe was scored for: (1) percentage of words produced without any sound errors; and (2) percentage of total consonants produced correctly, in the correct position within the target word.

Temporal-acoustic measures

Several temporal-acoustic measures were conducted on speech samples from selected probe sessions. These analyses were conducted in an attempt to better understand the perceptual findings of this investigation. Two baseline, one mid-treatment, and two end-of-treatment probes were chosen for analysis: baseline 3, baseline 5, probe 13, probe 21 and probe 22. One-half of the trained and untrained items from group 1 were randomly selected for temporal-acoustic measurements (see items marked in appendix 2).

All baseline and probe sessions were audiorecorded using a Sony ECM-T150 lapel microphone with either a Sony TC-D5M or Sony WM-D6C recorder. The recorded data (i.e. clinician's production, response interval and participant's production) were sampled at 10 kHz and stored using the Computerized Speech Lab (Kay Elemetrics 1994). Oscillographic and spectrographic displays were produced, linked, and utilized in all measurements.

Total word duration

Total word duration was chosen for measurement because treatment directly manipulated rate of production and an increase in word duration appeared to be likely. Total duration was measured from the onset of the first sound of the participant's production to the offset of the last sound. Specifically, the onset of voicing in nasals, vowels, liquids, and prevoiced stops and fricatives was determined by: (1) identifying the first cycle of vibration on an expanded oscillogram; (2) verifying the onset of vibration by identifying voicing bars and formants on the spectrogram; and (3) ascertaining that voicing did not occur prior to the selected onset point, by using the speak functions on CSL. Offset of voicing was determined by reversing the preceding process. For productions beginning with stops and fricatives that were not prevoiced, onset was determined by positioning the CSL cursor immediately preceding the release burst for stops and the noise onset for fricatives on the spectrographic display, with oscillographic and playback verification. Determining the offset of noise in fricatives and released stops was generally more difficult than determining the onset. In these cases, the spectrogram and oscillogram were used for determining cessation of fricated noise and the speak function was used for assistance in differentiating frication from background noise.

Response latency

The probe responses that served as the dependent measures were elicited in a modelling-repetition format, with no time constraints placed on the speaker's responses. Therefore, it was possible that the speaker may have adopted a strategy of responding slower, to allow for increased planning time, over the course of the investigation. Response latency was measured from the offset of the clinician's production to the onset of the participant's production using the onset/offset determination procedures described previously.

First syllable duration

Over the course of treatment, the speaker appeared to frequently exaggerate the length of the initial syllable of probe words. In order to attempt to confirm or deny this perceived initial syllable lengthening, measures of the duration of initial syllables were obtained. Durations were calculated from the onset of the first sound of the syllable to the offset of the last sound of the syllable.

Reliability

Transcription

Twenty percent of all baseline and probe sessions were randomly selected to calculate interjudge reliability. A second listener transcribed all items from the selected sessions. Point-to-point agreement between the original and second transcription was calculated for judgments of correctness of production of: (1) each word; and (2) each consonant. The percentage of agreement was 91% and 95% for words and consonants, respectively.

Temporal-acoustic measures

Twenty five percent of the words utilized in the original temporal-acoustic analyses were selected through stratified random sampling (i.e. equal numbers of items were selected across baseline and probe sessions) for remeasurement. The first author repeated all measurements at 4 weeks following the original analyses, using the same procedures. The percentages of measurements that fell within 10% (plus or minus) of the original measurements were as follows: (1) total word duration—88%; (2) response latency—80%; (3) first syllable duration—80%.

Differences in response latency values stemmed primarily from difficulties in determining the offsets of the clinician's utterances. Because a label microphone worn by the participant was used for recording, the clarity of the clinician's utterances was sometimes problematic, which appeared to account for the preceding difficulties. Most first syllable duration measurement discrepancies occurred in the determination of syllable offsets. The speaker typically did not segregate syllables during probes and coarticulatory influences served to make syllable offset determinations difficult.

Results

Perceptual Measures

Figure 1 displays the acquisition and response generalization findings relative to the percentage of words produced without any sound errors in probe sessions. The percentage of correct word productions is depicted for each group of words on separate graphs. Table 2 lists that percentage of correct consonants produced relative to the total number of consonants in the probe responses. These are grouped to correspond to each graph in figure 1.

As seen in figure 1, correct responding was relatively stable for all sound groups (with the exception of group 3C—/s/ blends) prior to the initiation of treatment. Following application of treatment to group 1 words, increases in correct production of trained and untrained group 1 words were seen. Specifically, the average baseline performance for trained group 1 items was 9% and 69% for correct words (figure 1) and correct consonants (table 2), respectively. In the last three probes during group 1 training, performance increased to an average of 72% for correct words and 91% for correct consonants.

Increases in correct productions of group 2, group 3B and group 3C words were also observed following treatment for group 1. Average baseline correct productions for words for each of these groups were as follows: (1) group 2—29%; (2) group 3B—16%; and (3) group 3C—28%. The average correct productions in the last three probes during group 1 treatment were as follows: (1) group 2—48%; (2) group 3B—35%; and (3) group 3C—60%. Corresponding changes in percentage of correct consonants were also observed from baseline to final probes during group 1 treatment: (1) group 2—77% (baseline) to 86% (end of group 1 treatment); (2) group 3B—69% to 79%; and (3) group 3C—66% to 92%.

Table 2. Percentage of consonants correct on probes

Probe	Group 1 3 syllables (1st syllable stressed) treated	Group 1 3 syllables (1st syllable stressed) untreated	Group 3A 3 syllables (3rd syllable stressed) treated	Group 3A 3 syllables (3rd syllable stressed) untreated	Group 2 3 syllables (2nd syllable stressed) untreated	Group 3B 4 syllables untreated	Group 3C /s/ blends untreated
BL 1	76	71	65	81	83	72	80
BL 2	72	76	70	76	79	67	74
BL 3	67	79	65	76	76	71	88
BL 4	71	73	70	71	77	67	82
BL 5	60	70	78	71	74	70	80
P 6	65*	71	65	71	70	66	92
P 7	84	79	61	67	76	65	68
P 8	87	78	78	67	80	71	92
P 9	88	76	78	62	82	79	92
P 10	87	79	57	67	82	73	94
P 11	84	84	74	71	82	66	90
P 12	85	88	70	81	83	74	92
P 13	89	81	65	81	82	80	90
P 14	78	82	74	76	83	81	96
P 15	87	83	74	86	86	76	92
P 16	91	89	74	71	81	75	90
P 17	91	84	65	76	88	84	92
P 18	89	92	74	76	86	80	88
P 19	89	88	70	90	88	84	87
P 20	93	91	74	76	89	81	94
P 21	90	89	70	81	83	81	91
P 22	91	88	70	76	86	76	92
P 23	87	88	83**	86	85	77	90
P 24	91	85	91	86	86	81	92
P 25	94	90	98	81	91	84	88
P 26	94	85	96	95	95	86	90
Follow-up	92	91	91	88	96	82	87

* First probe following initiation of treatment with group 1.
** First probe following initiation of treatment with group 3A.

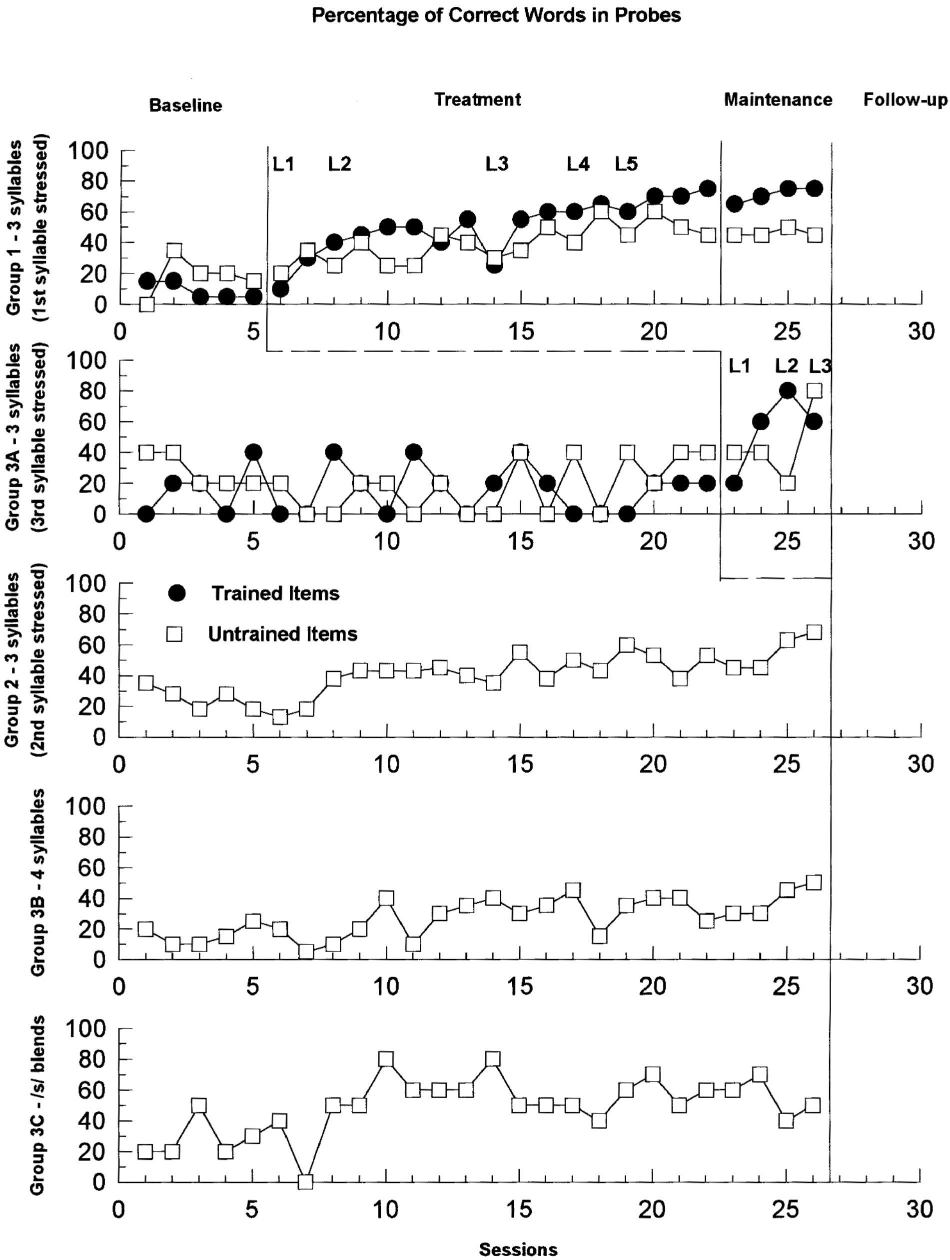

Figure 1. Percentage of correct word productions in probe sessions. Responses to trained items are represented by filled circles and responses to untrained items are represented by unfilled squares. The initiation of the different levels of treatment are indicated above the graphed data: The probe below each level indicator was conducted *after* that level was initiated.

Changes in group 3A productions were not apparent following treatment of group 1 words. Therefore, treatment was initiated with training items from group 3A. The 90% criterion was quickly reached with these items following initiation of treatment.

Response generalization to untreated items within groups was positive and

Table 3. Temporal-acoustic measures: group 1 items

Measure	Baseline 3	Baseline 5	Probe 13	Probe 21	Probe 22
Word duration (ms)					
mean	1389.4	1236.5	1305.9	1415.4	1299.4
sd	(315.2)	(325.1)	(327.9)	(485.4)	(434.9)
median	1324	1143	1233	1227	1180
25 %–75 %	(1214–1588)	(1063–1380)	(1050–1459)	(1088–1732)	(931–1543)
Response latency (ms)					
mean	516.9	585.2	652.9	549.3	410.4
sd	(147.3)	(378.0)	(213.2)	(281.7)	(256.0)
median	464	447	629	490	382
25 %–75 %	(431–546)	(382–630)	(526–815)	(323–621)	(226–473)
1st Syllable duration (ms)					
mean	578.3	473.8	574.9	686.3	614.0
sd	(150.6)	(175.4)	(173.5)	(282.9)	(225.6)
median	610	427	578	611	551
25 %–75 %	(457–682)	(339–583)	(448–670)	(481–754)	(455–748)
1st Syllable duration: word duration (ratio)					
mean	0.42	0.38	0.44	0.47	0.48
sd	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)
median	0.40	0.36	0.44	0.48	0.50
25 %–75 %	(0.37–0.48)	(0.32–0.45)	(0.38–0.55)	(0.40–0.51)	(0.39–0.54)

Table 4. Results from Friedman repeated measures analysis of variance on rank tests: comparisons across sampling times for the temporal-acoustic measures

Measure	Treated+untreated items	Treated items only	Untreated items
Word duration			
Chi-square	4.513	1.80	not conducted
P value	0.341	0.772	
Response latency			
Chi-square	9.108	12.457	not conducted
P value	0.058	0.014	
First syllable duration			
Chi-square	18.112	8.0	10.838
P value	0.001	0.092	0.028
First syllable: whole Word duration			
Chi-square	13.699	14.3	5.207
P value	0.008	0.006	0.267

strong for both group 1 and group 3A words. Follow-up at 3 weeks following treatment revealed no significant decline in gains achieved during treatment for all word groups.

Temporal-acoustic measures

Measures of central tendency and spread are reported in table 3 for all of the temporal-acoustic measures conducted on the group 1 items (10 treatment plus 10 non-treatment items). In addition to word duration, response latency, and first syllable duration, a calculation of the ratio of first syllable duration to total word duration was made.

Friedman Repeated Measures Analysis of Variance on Ranks tests were conducted to evaluate the differences in median values across sampling times. Analyses were conducted separately for: (1) word duration; (2) response latency; (3) first syllable duration; and (4) ratio of first syllable to whole word duration (table 4). For each of the durational measures, analyses were conducted for the entire group of 20 items (treatment plus non-treatment) and then the treated items alone.

The effect of sampling time was not significant for word duration or for response latency at $p < 0.01$ when considering either the entire group or only the treated items. There was a statistically significant difference (at $p < 0.01$) found in the first syllable duration median values among sampling times for the group of 20 items, but not for the treated or untreated items considered separately. Follow-up multiple comparison procedures (Dunn's method) revealed the median value of probe 5 (baseline) to be significantly lower than those of probes 13, 21 and 22. The median value of the other baseline probe (probe 3) was found to be not significantly different from any of the other probes. A statistically significant difference in median values among sampling times was also found for the ratio of first syllable duration to total word duration at $p < 0.01$ for the entire group of items and the treated items only. Follow-up comparison procedures indicated that the median ratio for probe 5 was significantly less than those for probes 13, 21 and 22 for the entire group of items. For the treated items, the median ratio for probe 5 was significantly less than those of probes 13 and 21. The median ratio for probe 3 was not significantly different from that of any of the other probes for both the entire group and the untreated items only.

Discussion

This investigation demonstrated that treatment resulted in improvements in sound production with an apraxic and aphasic speaker. Positive changes were observed not only in treated words, but in untreated words as well. These changes occurred despite the fact that this treatment did not directly target sound change (i.e. feedback was carefully controlled so that the speaker was never provided with information regarding the correctness of his word or sound productions).

As with any treatment that employs a combination of techniques, it is not possible to determine the relative contributions of the various treatment parameters to treatment effects. Repeated practice comprised a large portion of the treatment and there is the possibility that repetition alone might have resulted in the same changes. However, the positive response generalization effects (both to untrained exemplars of trained words and untrained words) speak against this speculation. That is, if repeated practice (or repeated exposure) accounted for treatment effects, positive increases in correct sound production should have been evident only for those words receiving treatment. Of course it may be argued that repeated practice resulted in a general improvement in facility of production of sequences of syllables, which could account for the generalization effects.

Similarly, it is not clear as to whether or not both the metronome and the hand-tapping were necessary. The use of both may appear to be redundant, but each was included to serve a different purpose. Hand-tapping necessitates an internally generated rhythm by the patient, although varying degrees of external control may be exerted on that rhythm. It was reasoned that internally generated rhythm may be more amenable to generalization than a completely externally generated rhythm. However, an externally generated rhythm was necessary to train and practice hand-tapping. The use of a metronome was selected over simple clinician modelling because it provided much more experimental control over rate and rhythm.

As mentioned previously, it was hoped that hand-tapping might serve as a means to mediate generalization from the training condition to nontraining conditions, such as probes. This patient was not instructed to attempt to generalize hand-tapping outside training and systematic measurements were not made of extratherapy hand-tapping. Initially, on-line notations of hand-tapping during probes were attempted, but were discontinued because of difficulties in observing these behaviours while performing on-line transcription (i.e. the patient often kept his hands on his lap under the therapy table). Obvious hand-tapping was not frequently observed during probes. In future investigations of such treatment, training patients to generalize tapping behaviour may be a worthwhile endeavour.

Although this treatment did appear to decrease the number of sound errors without any instruction on sound production, a significant number of sound and word errors remained. This speaker's remaining sound errors tended to occur as distortions of fricatives and affricates and word errors included those errors as well as some addition and substitution errors. It is not surprising that sound errors did remain following this treatment. As discussed by Square *et al.* (1997), the behaviours that characterize AOS include not only problems with the timing and/or seriation of gestures, but also with the generation of postures. This treatment was focused on temporal aspects of speech production and not on spatial aspects. Although rate/rhythm control treatments may positively impact the generation of articulatory postures (discussed below), it would not be expected that such a treatment would totally ameliorate sound errors that were derived from a 'degraded internalized schemata of the spatial coordinate system or inability to access the spatial coordinate system' (Square *et al*. 1997, p. 175). In order to eliminate remaining sound errors, this treatment could be relatively easily combined with sound production training. For example, feedback regarding sound accuracy could be provided along with feedback about timing and tapping accuracy.

Speculations as to why a rate and rhythm treatment may improve speech production in AOS have been offered previously (Dworkin and Abkarian 1996, Square and Martin 1994). One suggestion is that treatment such as that applied in this investigation may facilitate functioning of an internal oscillatory mechanism. Numerous researchers have provided support for the existence of central pattern generators (CPGs), or internal oscillatory mechanisms, that are important in the perception of time as well as in the control of timing of skilled movements (Keele *et al*. 1985, Rybak *et al*. 1997, Smith and Denny 1990, Treisman *et al*. 1994, Treisman *et al*. 1992). It is not clear as to whether the same mechanisms underlie perception and production, nor is it clear as to the number of oscillatory mechanisms present in the time-keeping system. Treisman *et al*. (1994) have proposed a calibrated temporal pacemaker model in which a central oscillator emits a relatively invariant oscillation that is resistant to perturbation, but that may be modulated by a calibration unit that flexibly adjusts the final output of pacemaker pulses. They also suggest the existence of a wide-spread 'system of parallel pacemakers' (p. 286) serving various temporal tasks.

It has been postulated that a pacemaker system underlies the temporal patterning of speech production (Gracco 1990), although the specification of such a system has not been elucidated. Whereas the articulators and respiratory system are clearly controlled by CPGs during biologic functions such as sucking and breathing, their neural control during speech relative to central pattern generators is not well understood (Smith and Denny 1990). Similarly, the role played by CPGs in AOS is not known. However, it is logical to hypothesize that the movement seriation problems that characterize AOS may stem from disruptions in a central oscillator. In considering Treisman *et al*.'s (1994) model, malfunctioning of the pacemaker system could occur at a number of levels that could impact movement sequencing: (1) the invariant oscillation provided by the central oscillator could be disrupted; (2) the calibration unit's responsiveness (to endogenous or exogenous stimuli) could be compromised; or (3) the pacemaker's integration with other pacemakers could be disturbed. The treatment used in this investigation was obviously a form of entrainment, or the phase-locking of movements (or other internal rhythms) to an external stimulus. In terms of pacemaker functioning, the entrainment may have served to reset the central oscillator or to strengthen the calibration unit's responsiveness to stimulation. Of course, any suggestion that this treatment facilitated CPG functioning is speculative. Future research concerning the periodicity of speech and nonspeech movements in AOS may provide insights about the role of CPGs in AOS, which may then allow the study of the effects of treatment on CPG functioning.

Related to timing control, but potentially distinct from the concept of CPGs, is the issue of articulatory slowing and its potential effect on articulation accuracy. Slowness of speech production is considered to be a core feature of AOS (Kent and Rosenbek 1983, McNeil *et al*. 1997). Kent and Rosenbek examined the acoustic patterns of the speech of seven AOS speakers and reported two forms of slowed rate during phrase, sentence, and multisyllabic word production, which they termed articulatory prolongation and syllable segregation. The subject in this investigation exhibited both types of prolongation during production of experimental words in training with the metronome as well as in probes. Although overall word durations in probes did not change significantly from baseline to final sessions (i.e. the subject did not appear to have slowed his overall rate of production during probes), the slowed rate of production during treatment may have had a facilitory effect on articulation that generalized to probes. Specifically, it has been suggested that slowed rate may heighten sensory feedback during articulation and may provide additional motor planning time (Square *et al*. 1997).

Results of the response latency analysis suggest that the subject did not significantly modify his response times in probes across the course of the investigation to allow for additional motor planning time. Conversely, the increase in first syllable durations and ratios of first syllable to total durations may reflect the subject's attempts to provide enhanced feedback and/or added planning time during word production. However, because the preceding findings did not apply with both analysed baseline probes, the first syllable duration results should be considered to be inconclusive. As indicated above, lack of durational changes on probes need not be interpreted as indicating that slowing of rate in treatment was not important. The gains in enhanced feedback and motor planning time that were possibly achieved in treatment may have been sufficient to re-establish more effective and efficient motor planning that was then realized in probes without the need for reduced rate. As indicated previously, it is not possible to determine which components of this treatment package were critical for effecting positive changes in consonant production. One may consider the observation that increases in consonant accuracy continued during rate increases (level 4) to be antithetical to the concept that slowed rate facilitated consonant production. However, such effects may not have occurred if level 4 had not been preceded by levels 1–3, and level 4 rate was still a slowed rate with respect to the patient's initial rate.

Of course, increased motor planning effectiveness can only be inferred through observation of increased articulatory accuracy in this investigation. However, in future investigations, the use of a response priming paradigm, such as that used by Rogers and Storkel (1999), may reveal changes in processing efficiency following treatment.

Other directions for future research in this area should include an examination of the effects of repetition versus use of the metronome plus repetition. Because of the positive response generalization observed in this investigation, results from an alternating treatments design would be difficult to interpret. However, the current design could be modified to include an exposure control group of items that would receive as much repetition practice as items that receive the metronome plus repetition treatments. Results from the exposure control items could then be compared to another group of items that would receive no treatment and would be probed only during baseline and following treatment (to avoid repeated practice in probes), as well as to the metronome plus repetition items.

Evaluation of the effects of use of the metronome at a slowed rate versus at the speaker's habitual rate could also be potentially useful. The metronome rate selected for this investigation was approximately 50% slower than the speaker's typical rate of production of multisyllabic words. There is no means of determining from this investigation whether or not this was an optimal rate for training. There has been evidence to suggest that CPGs may have a limited range of sensitivity to entrainment (Treisman *et al*. 1994) (i.e. the rate of oscillation used for entrainment typically must be relatively close to the natural rate of oscillation) and the rate of oscillation used in this investigation may have not been optimal. The evaluation of rate effects may be amenable to study through the use of a reversal design in which rate is manipulated and effect on articulation is measured.

This study represents an initial investigation of the effects of rate and rhythm training on perceived articulation as determined through broad phonetic transcription. It was not designed to evaluate the clinical efficacy of this type of treatment. Issues such as optimal length of treatment, stimulus generalization effects, and functional outcomes of treatment should be examined in future investigations before this treatment is advocated for clinical use. Additionally, replications of this investigation are obviously necessary for the purposes of external validity. However, this type of treatment does appear to have promise for the treatment of sound errors in AOS and aphasia and certainly warrants further study.

Acknowledgements

This research was supported by the Department of Veterans Affairs, Rehabilitation Research and Development. This investigation was presented as a poster at the annual Clinical Aphasiology Conference, Asheville, NC, June 1998.

References

Albert, M. L., Spades, R. W. and Helm, N. A., 1973, Melodic intonation therapy for aphasia. *Archives of Neurology*, **29**, 130–131.

Dabul, B., 1979, *Apraxia Battery for Adults* (Tigard, OR: C.C. Publications).

Duffy, J. R., 1995, *Motor Speech Disorders Substrates, Differential Diagnosis, and Management*. St. Louis, MO: Mosby-Year Book, Inc.

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

Dworkin, J. P. and 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.

Fromm, D., Abbs, J. H., McNeil, M. R. and Rosenbek, J. C., 1982, Simultaneous perceptual-physiological method for studying apraxia of speech. *Clinical Aphasiology*, **10**, 155–171.

German, D. J., 1990, *Test of Adolescent/Adult Word Finding*. Allen, Texas: DLM.

Gracco, V. L., 1990, Characteristics of speech as a motor control system. In G. Hammond (Ed.), *Cerebral Control of Speech and Limb Movements: Advances in Psychology* (pp. 3–28). New York: Elsevier.

Hardcastle, W. J, 1987, Electropalatographic study of articulation disorders in verbal dyspraxia. In J. H. Ryals (Ed.), *Phonetic Approaches to Speech Production in Aphasia and Related Disorders* (pp. 113–136). Boston: College-Hill Press.

Helm, N., 1979, Management of palilalia with a pacing board. *Journal of Speech and Hearing Disorders*, **44**, 350–353.

Itoh, M., Sasanuma, S., Hirose, H., Yoshioka, H. and Ushijima, T., 1980, Abnormal articulatory dynamics in a patient with apraxia of speech: X-ray microbeam observations. *Brain and Language*, **11**, 66–75.

Kay Elemetrics Corporation, 1994, *Computerized Speech Lab* (Lincoln Park, NJ: Author).

Keele, S. W., Pokorny, R. A., Corcos, D. M. and Ivry, R., 1985, Do perception and motor production share common timing mechanisms: a correlational analysis. *Acta Psychologica*, **60**, 173–191.

Kent, R. D. and Rosenbek, J. C., 1983, Acoustic patterns of apraxia of speech. *Journal of Speech and Hearing Research*, **26**, 231–249.

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

McNeil, M. R., Robin, D. A. and 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.

Naeser, M. A. and Helm-Estabrooks, N., 1985, CT scan lesion localization and response to melodic intonation therapy with nonfluent aphasia cases. *Cortex*, **21**, 203–223.

Porch, B., 1967, *Porch Index of Communicative Ability*. Palo Alto, CA: Consulting Psychologists Press.

Porch, B., 1981, *Porch Index of Communicative Ability*. Palo Alto, CA: Consulting Psychologists Press.

Rogers, M. A. and Storkel, H. L., 1999, Planning speech on syllable at a time: the reduced buffer capacity hypothesis in apraxia of speech. *Aphasiology*.

Rubow, R. T., Rosenbek, J. C. and Collins, M. J., 1982, Vibrotactile stimulation for intersystemic reorganization in the treatment of apraxia of speech. *Archives of Physical Medicine and Rehabilitation*, **63**, 150–153.

Rybak, I. A., Paton, J. F. R. and Schwaber, J. S., 1997, Modelling neural mechanisms for genesis of respiratory rhythm and pattern. II. Network models of the central respiratory pattern generator. *Journal of Neurophysiology*, **77**, 2007–2026.

Saffran, E. M., Berndt, R. S. and Schwartz, M. F., 1989, The quantitative analysis of agrammatic production: procedure and data. *Brain and Language*, **37**, 440–479.

Shane, H. C. and Darley, F. L., 1978, The effect of auditory rhythmic stimulation on articulatory accuracy in apraxia of speech. *Cortex*, **14**, 444–450.

Simmons, N. N., 1978, Finger counting as an intersystemic reorganizer in apraxia of speech. In R. H. Brookshire (Ed.), *Clinical Aphasiology: Conference Proceedings* (pp. 174–179). Minneapolis, MN: BRK.

Smith, A. and Denny, M., 1990, High-frequency oscillations as indicators of neural control mechanisms in human respiration, mastication, and speech. *Journal of Neurophysiology*, **63**, 645–758.

Southwood, H., 1987, The use of prolonged speech in the treatment of apraxia of speech. In R. H. Brookshire (Ed.), *Clinical Aphasiology: Conference Proceedings* (pp. 277–287). Minneapolis, MN: BRK.

Sparks, R. W. and Deck, J. W., 1994, Melodic intonation therapy. In R. Chapey (Ed.), *Language Intervention Strategies in Adult Aphasia* (pp. 368–386). Baltimore, MD: Williams & Wilkins.

Square, P. A. and Martin, R. E., 1994, The nature and treatment of neuromotor speech disorders in aphasia. In R. Chapey (Ed.), *Language Intervention Strategies in Adult Aphasia* (third edition, pp. 467–499). Baltimore, MD: Williams & Wilkins.

Square, P. A., Roy, E. A. and Martin, R. E., 1997, Apraxia of speech: Another form of praxis disruption. In L. J. Gonzalez-Rothi and K. M. Heilman (Eds.), *Apraxia: The Neuropsychology of Action* (pp. 173–206). (Hove, UK: Psychology Press).

Treisman, M., Cook, N., Naish, P. L. N. and MacCrone, J. K., 1994, The internal clock: electroencephalographic evidence for oscillatory processes underlying time perception. *The Quarterly Journal of Experimental Psychology,* **47A**, 241–289.

Treisman, M., Faulkner, A. and Naish, P. L. N., 1992, On the relation between time perception and the timing of motor action: evidence for a temporal oscillator controlling the timing of movement. *The Quarterly Journal of Experimental Psychology*, **45A**, 235–263.

Wambaugh, J. L. and Doyle, P. J., 1994, Treatment for acquired apraxia of speech: a review of efficacy reports. *Clinical Aphasiology*, **22**, 231–243.

Wertz, R. T., LaPointe, L. L. and Rosenbek, J. C., 1984, Apraxia of speech in adults: the disorder and its management. Orlando, FL: Grune & Stratton, Inc.

Yorkston, K. M. and Beukelman, D. R., 1981, *Assessment of Intelligibility of Dysarthric Speech*. Austin, TX: PRO-ED.

Appendix 1: examples of misarticulations

abandon → dəbændən
antelope → ætɪlop
badminton → bædbɪtn
beautiful → dəbjutəful
benefit → bɛnəbɪt
camisole → kæbɪsol
casserole → kæswol
chivalry → ʃɪbərɪ
computer → ʌmpjutɚ
dignity → dɪgnənɪ
estimate → ɛtɪmet
ferocious → dʒərosəs
habitat → hæɪtæt
restaurant → drɛsəwɑnt
galaxy → kæləsi
therapy → θɛrbəpi
moccasin → mɑksɪsɪn
November → bobɛmbɚ
treasury → trɛdʒərɪ
replica → drɛklɪkə
recipe → dwɛsɪpi
Winchester → dwɪntʃɛtɚ
satisfy → sætɪfpaɪ
September → sɛtembɚ
umbrella → bʌbrɛdə

Appendix 2: experimental stimuli

Group 1	Group 2	Group 3A
stress on first syllable	stress on second syllable	stress on third syllable
treated items		*treated items*
antelope	abandon	absolute
badminton	aquatic	columnade
benefit	carnation	diagnose
camisole	companion	represent
casserole*	complexion	serenade
celibate	Dalmation	*response generalization items*
dangerous*	destruction	cavalier
estimate*	dynamic	chandelier
fabulous	efficient	guarantee
galaxy*	ferocious	potpourri
generate*	gorilla	seventeen
habitat	hydraulic	
industry	important	Group 3B
recipe*	inherit	four syllable words
regular*	mortician	acupuncture
restaurant	November	aluminum
satisfy*	proficient	architecture
saxophone*	spaghetti	deodorant
therapy*	tomato	elevator

wonderful

response generalization items

beautiful
calcium*
casual
chivalry
corduroy
cranium
crocodile
dignity*
envelope*
extrovert
hospital*
ignorant
management*
moccasin
replica*
rosary*
sensitive*
Subaru*
treasury*
Winchester

umbrella
acoustic
cadaver
cirrhosis
Columbus
computer
December
embarrass
fanatic
flamingo
October
petunia
piano
provision
relinquish
remember
salvation
September
tornado
tradition
transparent

embolism
January
secretary
superior
thermometer
barometer
catastrophe
category
equivalent
February
filibuster
heredity
television
unusual

Group 3C

/s/ cluster words

spread
sprinkle
spruce
straighten
strawberry
sprain
sprocket
stranger
strangle
stricken

* items selected for temporal-acoustic measurements

Appendix 3: treatment

Level one—clinician model, unison production, beginning patient production

A. S/T[a]
B. Metronome setting: 93 bpm
C. Treatment steps:
   1. CM[b]—1 production
   2. Patient taps along (no verbal production), while clinician produces word—5 productions
   3. UPT[c]—3 productions
   4. PPT[d]—1 production
D. Target item presentation: Clinician presents experimental words (treatment items only) in random order. Clinician presents as many items per session as possible.
E. Feedback: Clinician provides positive or negative feedback about tapping, production of correct number of syllables, and/or production of syllables on beat, *but not about sound accuracy*
F. Scoring: "+" or "−" for PPT step ("+" = correct use of tapping with production of correct number of syllables on the beat)
G. Criterion: 95% accuracy for entire treatment session in two consecutive sessions

Level two—faded clinician model, repeated patient production

A. S/T
B. Metronome setting: 93 bpm
C. Treatment Steps:
   1. CM—1 production
   2. PPT—3 productions
      a) If any errors in tapping to the beat or in producing correct number of syllables—CM (1 production) plus UPT (3 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: same as Level 1
F. Scoring: "+" or "−" for first production of PPT step
G. Criterion: 95% accuracy for entire treatment session in two consecutive sessions

Level three—no clinician model, repeated patient production

A. S/T
B. Metronome setting: 93 bpm
C. Treatment steps:
   1. Clinician says word with normal rate and prosody (no metronome or tapping)
   2. PPT—3 productions
      a) If any errors in tapping to the beat or in producing correct number of syllables—CM (1 production) plus PPT (3 productions); if errors remain—CM (1 production) plus UPT (3 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: same as Level 1
F. Scoring: "+" or "−" for first production of PPT step
G. Criterion: 95% accuracy for entire treatment session in two consecutive sessions

Level four—increased rate of production

A. Metronome setting: 100–110 bpm
B. Treatment Steps: same as Level Three
C. Target item presentation: same as Level 1
D. Feedback: same as Level 1
E. Scoring: "+" or "−" for first production of PPT step

Criterion: 95% accuracy for entire treatment session in two consecutive sessions and no decrease in accuracy of trained items on probes

Step five—syncopated production

A. Clinician explains concept of syncopation: The target word will be produced in two beats, with the first syllable on downbeat of handtap, the second syllable on upbeat, and third syllable on the downbeat of second handtap.
B. S/T
C. Metronome setting: 100 bpm
D. Treatment steps: apply syncopation sequentially at each of preceding levels, beginning with Level One

[a] *Schematic/tapping review* (*S/T*): Clinician explains/reviews schematic of stress pattern for words under treatment. Clinician and patient practice tapping with the schematic (__ __ ↓).
[b] *Clinician model* (*CM*): Clinician produces word with metronome (one syllable per beat), while tapping
[c] *Unison production and tapping* (*UPT*): Clinician and patient simultaneously produce target word, while tapping
[d] *Patient production and tapping* (*PPT*): Patient produces target word, while tapping. Clinician provides no model or assistance.

Note: Levels 4 and 5 were included at that subject's request in an attempt to naturalize speech production during the treatment procedure. Accuracy of consonant production on probes was evaluated as part of the level 4 criterion to ensure that increased rate had not detrimentally affected achieved gains in consonant accuracy prior to attempting syncopation.