The Effect of Visual Cues Provided by Computerized Aphasia Treatment

Research in the past decades has demonstrated that computerized aphasia treatment can yield positive outcomes in verbal expression (Choe, Azuma, Mathy, Liss, & Edgar, 2007), auditory discrimination (Mills, 1982), reading comprehension (Petheram, 1996), and writing (Deloche, Dordain & Kremin, 1993). In the current healthcare environment where long-term rehabilitation is rarely supported by the third-party payers, computer programs can offer therapy that is affordable, accessible, and beneficial. Although evidence is being accumulated to support the effectiveness of computerized aphasia treatment, it is largely unknown which variables are crucial in enhancing the outcome of such treatments. Given that many computer-based treatment programs are commercially available and, more importantly, individuals with aphasia and their caregivers are increasingly interested in purchasing those products, it is imperative to identify the critical components in a successful treatment program.

Of many potentially influential factors, the current study focused on the type of cueing and compared two cueing conditions: auditory-visual cues (i.e., listening to speech while looking at the speaker's face) vs. auditory-only cues (i.e., listening to speech without the speaker's face shown). In aphasia and apraxia rehabilitation, visual cueing has been considered to play a crucial role in improving verbal production (Rosenbeck, 1978; Wambaugh, Kalinyak-Fliszar, West, & Doyle, 1998). Thus, patients are routinely encouraged to look at their clinicians when they are practicing word retrieval and verbal production during traditional face-to-face speech therapy sessions. More recently, computer treatment programs are providing visual cues such as video clips zooming into a speaker's mouth producing target words (Bungalow Software Inc) or an animated talking head that interacts with patients (Cherney, Halper, Holland, & Cole, 2008). Despite the frequent use of visual cues in speech therapy, it has not been examined: (1) whether individuals with aphasia and verbal apraxia improve their auditory comprehension by looking at a speaker's face; and (2) whether the visual input provided during speech therapy enhances verbal expression. The current study attempted to answer those two questions.

Methods

Participants

Two individuals with chronic aphasia and verbal apraxia participated in the study. Table 1 summarizes their demographic profiles. The first participant, TV, was a 55-year-old male who had suffered from a stroke 14 years ago. He received an associate degree from a community college and formerly worked as a machinist at a sheet metal company. He spoke short phrases and sentences, and wrote single words and numbers. The second participant, ML, was a 56-year-old female who had suffered from a stroke 4 years ago. She had a college degree and previously taught mathematics at a charter school. She spoke fluently with occasional word-finding difficulties. Both participants demonstrated auditory and visual acuity within functional limits for the experimental procedure of the study. During the pre-treatment phase, the *Western Aphasia Battery – Revised* (Kertesz, 2007), the *Cognitive Linguistic Quick Test (CLQT)* (Helm-Estabrooks, 2001), and the *Apraxia Battery for Adults – Second Edition* (Dabul, 2000) were administered. Table 2 provides an overview of the test results. TV presented with moderate Broca's aphasia, mild-to-moderate apraxia, and moderate cognitive-linguistic deficits. ML demonstrated mild anomic aphasia and mild verbal apraxia. Her scores in all the domains of the *CLQT* were within normal limits.

Stimuli

Stimuli for auditory-visual comprehension tasks were selected from the Auditory Comprehension section of the *Boston Diagnostic Aphasia Examination – Third Edition (BDAE-3)* (Goodglass, Kaplan, & Barresi, 2001) and the *Revised Token Test (RTT)* (McNeil & Prescott, 1978). Forty-two single word items from the *BDAE-3* were evenly divided into three conditions: auditory-visual, auditory-only, and visual-only. Twenty yes/no questions from the *BDAE-3* and forty sentences from the *RTT* were divided into two conditions: auditory-visual, and auditory-only.

The primary task for the computer practice was confrontation naming. To make the task functionally relevant, word stimuli were selected by the participants. Participants' ability to name target pictures prior to the treatment phase was probed at the baseline assessment. Each verbal response was scored using a modification of the *Porch Index of Communicative Ability* scoring system (Porch, 1981). Table 3 summarizes the 16-point scoring system. In addition, to control the motivation and motor demands of the selected words, an importance rating of the target words and the articulatory demand on consonant production were obtained. Based on the initial assessments, thirty target items were divided into three practice conditions: auditory-visual practice, auditory-only practice, and no practice (control). Twenty words in practice conditions were programmed for computer practice by using Microsoft PowerPoint software. Each target word was presented on six consecutive slides with an increasing level of support. Appendices A and B present examples of practice slides. The program for each practice condition lasted for approximately 15 minutes.

Procedure

TV attended a 30-minute practice session five times a week for four weeks. ML attended ten 30-minute practice sessions over a four-week period. In each practice session, the participants' performances on naming target items were closely monitored. TV's progress on twenty practice words and ten control words was probed at the end of the four-week period (post-treatment), after four weeks of no practice (maintenance), and after four months of no practice (follow-up). ML was administered the post-treatment and maintenance assessments. Her follow-up assessment is scheduled in early May.

Results

Auditory-Visual Comprehension

Table 4 presents the two participants' accuracy scores on the single words and yes/no questions selected from the *BDAE-3*. TV's single word comprehension scores were the same in the auditory-visual and the auditory-only conditions. He had noticeable difficulty understanding single words that were presented only visually with no sounds. In contrast, TV gave more correct answers to yes/no questions in the auditory-visual than the auditory-only condition. ML missed only one item, suggesting ceiling effect. Table 5 summarizes the results on items from the *RTT*. TV had slightly higher mean scores in the auditory-visual than in the auditory-only condition, whereas ML showed a reverse trend. However, Wilcoxon signed-ranks tests indicated that the mean differences were not significant for TV ($p=.148$) and ML ($p=.178$).

Verbal Naming

Figures 1 and 2 show how much support TV and ML, respectively, needed from the computer program to say target words in each session. TV required gradually decreasing amounts of support in both practice conditions. In contrast, ML needed noticeably less support for the words in the auditory-visual than in the auditory-only condition from Sessions 4 through 9. Figures 3 and 4 respectively present TV's and ML's progress on thirty target words at the assessment sessions. Table 6 summarizes the results of Wilcoxon signed-ranks tests on the data.

Significant improvements were more consistently observed in the auditory-visual than in the auditory-only condition. Only ML showed a generalization to untrained items. Table 7 presents the number of words each participant spontaneously produced without any cues or modeling provided by the clinicians. The progress observed in the two practice conditions was comparable. Both participants demonstrated generalizations in this measure.

Discussion

Two individuals with chronic aphasia and verbal apraxia practiced naming functionally relevant words either with auditory-visual or auditory-only cues provided by a computer program. The participants demonstrated more rapid and consistent improvements in the auditory-visual than in the auditory-only condition. Further investigation will elucidate the association, or lack thereof, between visual processing skills and the ability to utilize visual cues for speech practice.

Refereces

Cherney, L.R., Halper, A.S., Holland, A.L., & Cole, R. (2008). Computerized script training for aphasia: preliminary results. *American Journal of Speech-Language Pathology*, 17, 19-34.

Choe, Y., Azuma, T., Mathy, P., Liss, J.M., & Edgar, J. (2007). The effect of home computer practice on naming in individuals with non-fluent aphasia and verbal apraxia. *Journal of Medical Speech-Language Pathology*, 15 (4), 407-421.

Dabul, B. (2000). *Apraxia Battery of Adults – Second Edition*. Austin, TX: Pro-Ed.

Deloche, G., Dordain, M., & Kremin, H. (1993). Rehabilitation of confrontation naming in aphasia: relations between oral and written modalities. *Aphasiology*, 7 (2), 201-216.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *Boston Diagnostic Aphasia Examination – Third Edition.* Philadelphia, PA: Lippincott Williams & Wilkins.

Helm-Estabrooks, N. (2001). *Cognitive Linguistic Quick Test*. San Antonio, TX: Harcourt Assessment Company.

Kertesz, A. (2007). *Western Aphasia Battery – Revised*. San Antonio, TX: Harcourt Assessment Company.

Mills, R.H. (1982). Microcomputerized auditory comprehension training. In R.H.Brookshire (Ed.) *Clinical Aphasiology Conference Proceedings* (BRK Publishers, Minneapolis).

Petheram, B. (1996). Exploring the home-based use of microcomputers in aphasia therapy. *Aphasiology*, 10 (3), 267-282.

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

Rosenbeck, J. C. (1978). Treating apraxia of speech. In D.F. Johns (Ed.) *Clinical Management of Neurogenic Communicative Disorders* (pp.191-241). Boston: Little, Brown.

Wambaugh, J.L., Kalinyak-Fliszar, M.M., West, J.E., & Doyle, P.J. (1998). Effects of treatment for sound errors in apraxia of speech and aphasia. *Journal of Speech, Language, and Hearing Research*, 41, 725-743.

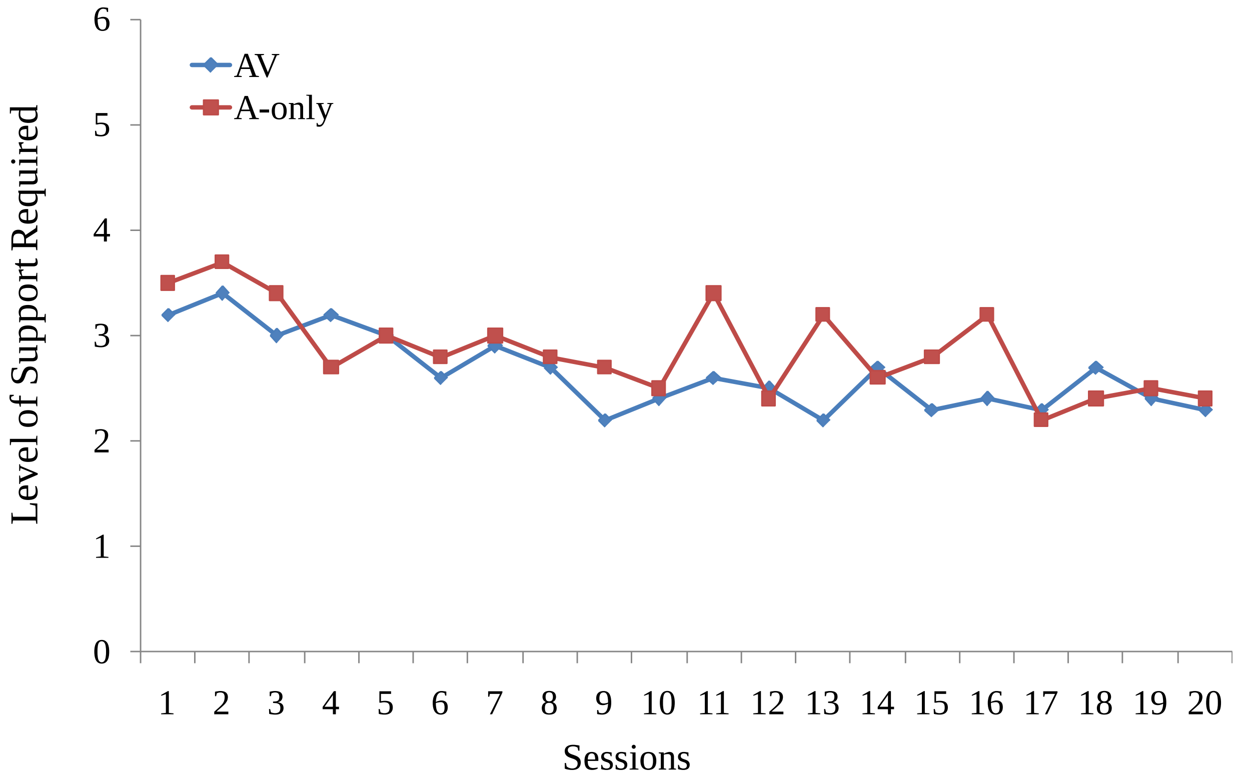


Figure 1. Average Level of Support Required for TV to Produce Target Words in Each Practice Session

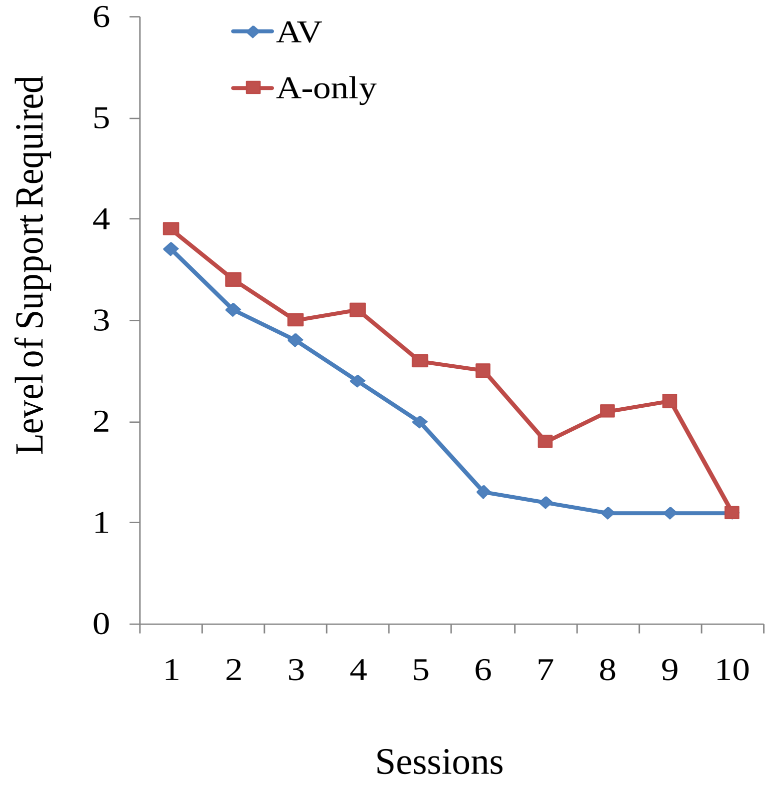


Figure 2. Average Level of Support Required for ML to Produce Target Words in Each Practice Session

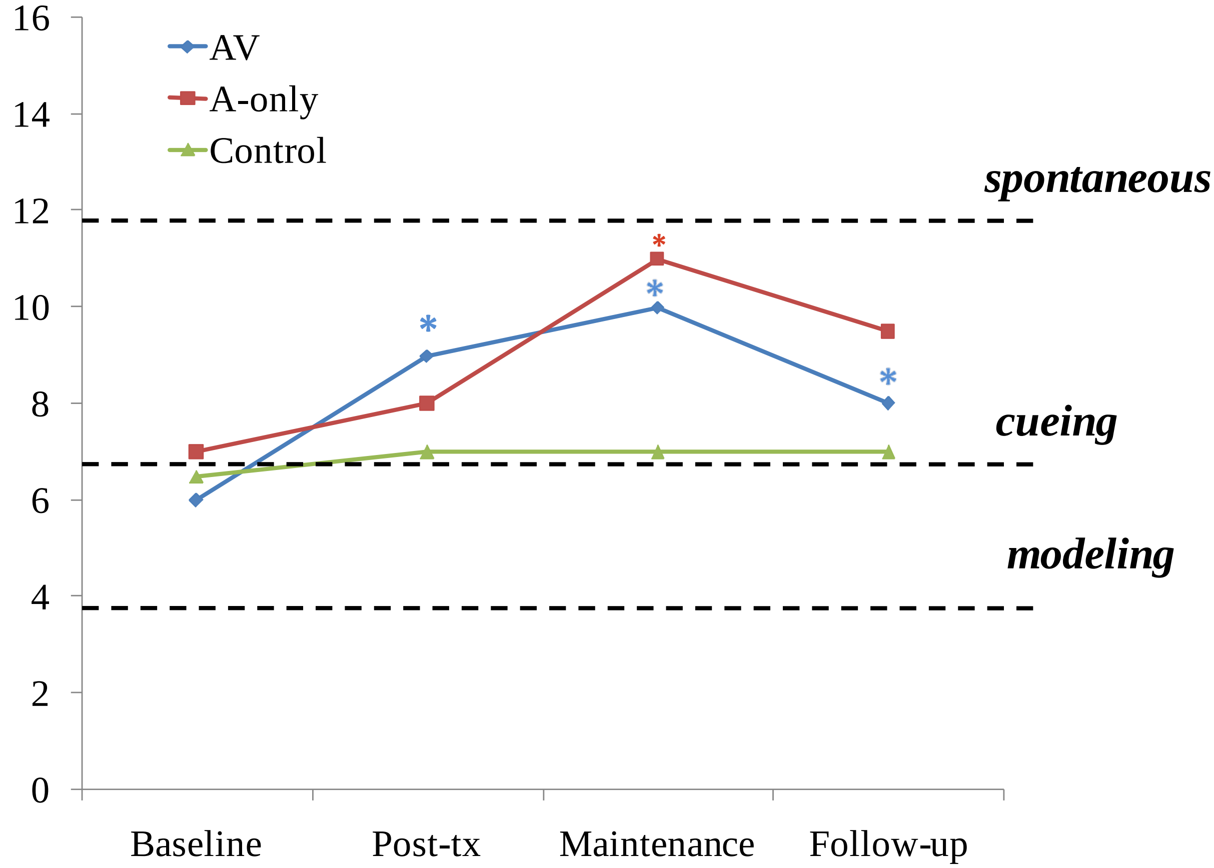


Figure 3. Median Naming Scores for the Auditory-Visual Practice, Auditory-only Practice, and Control Conditions for Participant TV

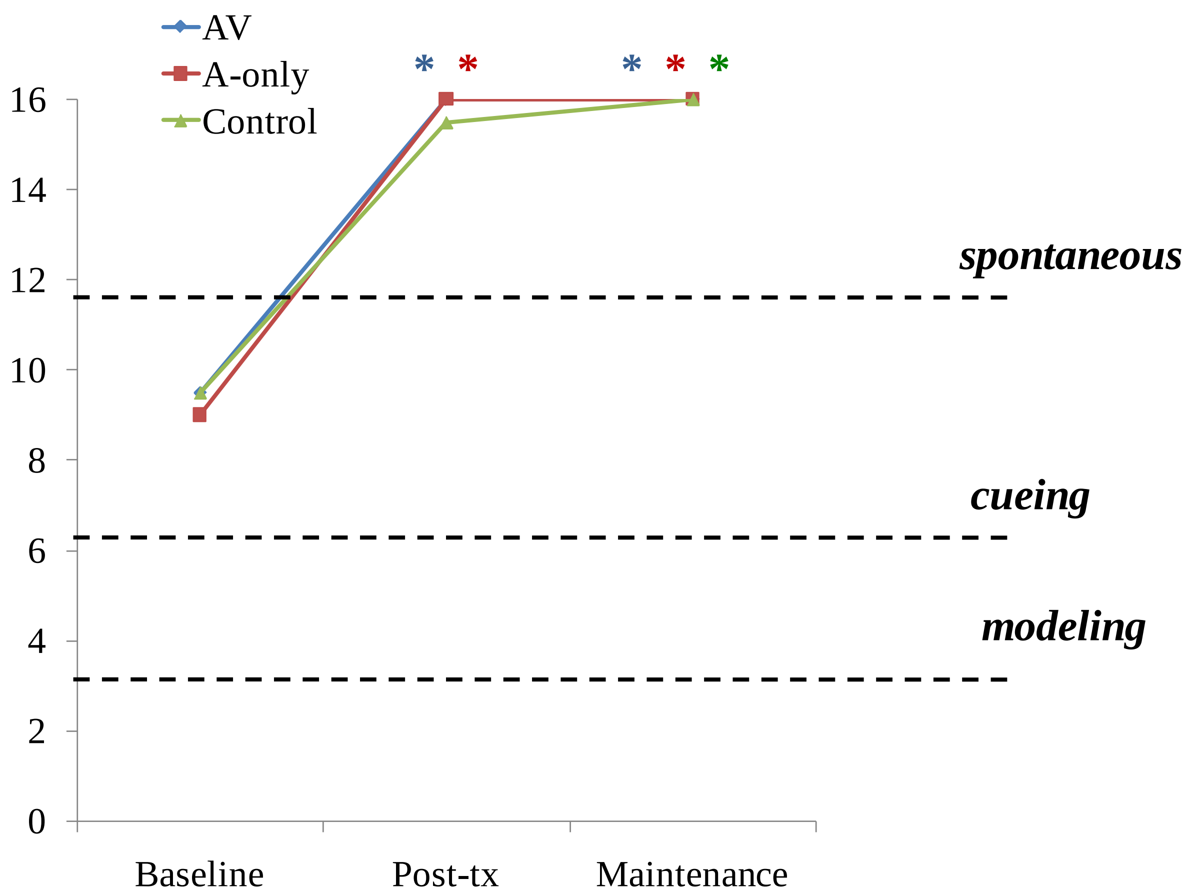


Figure 4. Median Naming Scores for the Auditory-Visual Practice, Auditory-only Practice, and Control Conditions for Participant ML

Table 1
Demographic Characteristics of the Study Participants

Participant	Etiology	Gender	Age	Time Postonset (yrs; mos)	Years of Education	Former Occupation
TV	CVA	M	55	14;5	14	Machinist
ML	CVA	F	56	4;1	16	Teacher

Table 2
Overview of the Standardized Test Results

Participant	*Western Aphasia Battery – Revised*		*Apraxia Battery for Adults – Second Edition*		*Cognitive Linguistic Quick Test*
	Aphasia Quotient	Aphasia Type	Verbal apraxia	Limb/Oral apraxia	Composite Severity Rating
TV	53.5	Broca's	Mild – Moderate	Mild	Moderate
ML	94.9	Anomic	None – Mild	None	WNL

Table 3

Hierarchical Scoring System for Naming Responses

Score	Response Type	Category	Description
16	Complete	Spontaneous	Correctly says the target word without any cue within three seconds.
15	Delayed	Spontaneous	Correctly says the target word without any support but after more than three seconds' delay.
14	Phonemic Error	Spontaneous	Incorrect phonemes are pronounced but spontaneously corrected (e.g., "tat……cat" for a target 'cat').
13	Delayed Phonemic Error	Spontaneous	After more than three seconds' delay, incorrect phonemes are pronounced but spontaneously corrected.
12	Self-Corrected	Spontaneous	Responds with a wrong word and then self-corrects (e.g., "dog….cat" for a target 'cat').
11	Semantic Cue	Cueing	Correctly says the target word after a phrase or sentence providing a semantic cue (e.g., "it meows" for the target 'cat').
10	Word Shape Cue	Cueing	Correctly says the target word when an initial letter and total number of letters are given in a written form (e.g., 'C_ _' for a target 'cat').
9	Whole Word Written Cue	Cueing	Correctly says the target word when the whole word is presented in a written form.
8	Initial Sound Cue	Cueing	Requires initial sound cue (e.g., the /k/-sound for a target 'cat') before correctly producing the target word. The cue can be repeated once on request.
7	Lip Shape Cue	Cueing	Requires seeing a clinician silently mouth the word before correctly producing the target word. The cue can be repeated once on request.
6	Whole Word Spoken Cue	Modeling	Correctly says the target word after it has been spoken by the clinician. Modeling may be repeated once by request.
5	Repeated Presentation	Modeling	Correctly says the target word after watching the clinician repeat the word five times.
4	Simultaneous Production	Modeling	Correctly produces the target word during five times of in unison repetitions with the clinician.
3	Tactile Cue	Tactile	Correctly produces a target word with touch cues in conjunction with in unison repetitions from the clinician.
2	Incomplete	Not Produced	Produces an approximation but cannot completely produce the word.
1	Incorrect	Not Produced	Produces none of the phonemes in a target words.
0	No Response	Not Produced	Produces no response or unrelated response (e.g., stereotypic utterance).

Table 4

Accuracy in Auditory Comprehension on Stimuli from the BDAE-3 (%)

Participants	Single Words			Yes/No questions	
	Auditory-Visual	Auditory-only	Visual-only	Auditory-Visual	Auditory-only
TV	71.2	71.2	21.4	90	60
ML	100	100	92.9	100	100

Table 5
Mean Communication Scores on Stimuli from the RTT (/15)

Participants	Subtest 1		Subtest 2		Subtest 3		Subtest 4	
	A – V	A – only	A – V	A – only	A – V	A – only	A – V	A – only
TV	12.7	11.5	12.4	10.1	10.1	9	DNC*	DNC*
ML	15	15	14.2	14.8	14.2	14.4	14.4	14.3

AV: Auditory-Visual condition
A-only: Auditory-only condition

* TV did not complete this subtest.

Table 6

Results of Wilcoxon Signed Tests Comparing Post-treatment, Maintenance, and Follow-up Assessments to Baseline Performance for Participants TV and ML

Practice Condition	Post-tx		Maintenance		Follow-up
	TV	ML	TV	ML	TV
AV Practice	2.532 p = .011*	2.889 p =.004*	2.494 p = .013*	2.879 p =.004*	2.506 p = .012*
A-only Practice	2.043 p = .041	2.642 p =.008*	2.829 p = .005*	2.879 p =.004*	2.383 p = .017
Control	0.552 p = .581	2.392 p =.017*	2.032 p = .042	2.549 p =.011*	1.980 p = .048

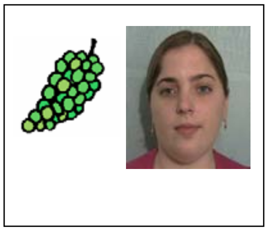
Note: * indicates significant at $p < .0167$ (Bonferroni corrected alpha) for TV and $p < .025$ for ML

Table 7

Number of Words that the Participants Produced without Any Support from the Clinicians at Each Assessment

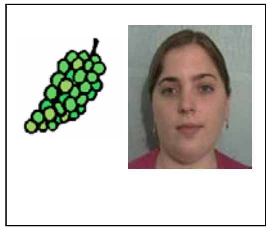
Practice Condition	Baseline		Post-tx		Maintenance		Follow-up
	TV	ML	TV	ML	TV	ML	TV
AV Practice	0	0	1	10	2	10	3
A-only Practice	0	0	2	7	4	10	4
Control	0	0	0	7	1	8	2

APPENDIX A
EXAMPLE OF SLIDES IN AUDITORY-VISUAL CONDITION



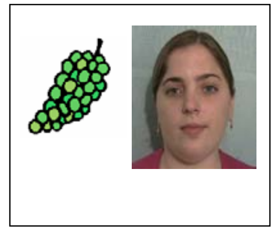
Auditory cue: What is this?

[Slide 1]



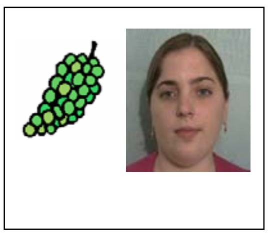
Auditory cue: You make wine with this fruit. What is this?

[Slide 2]



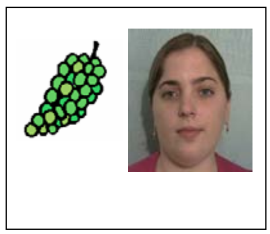
Auditory cue: It starts with /gr/. What is this?

[Slide 3]



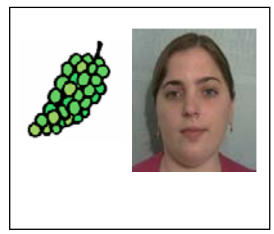
Auditory cue: It's grapes. Say grapes.

[Slide 4]



Auditory cue: Listen to me say it. Grapes (x5). Say grapes.

[Slide 5]

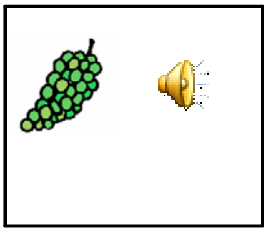


Auditory cue: Now say it with me. Grapes (x5). What is this?

[Slide 6]

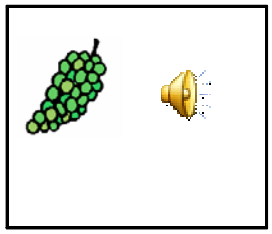
APPENDIX B

EXAMPLE OF SLIDES IN AUDITORY-ONLY CONDITION



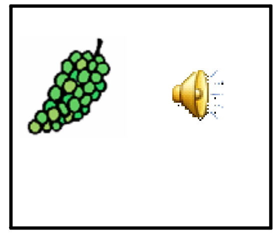
Auditory cue: What is this?

[Slide 1]



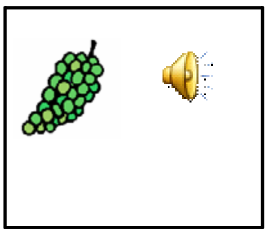
Auditory cue: You make wine with this fruit. What is this?

[Slide 2]



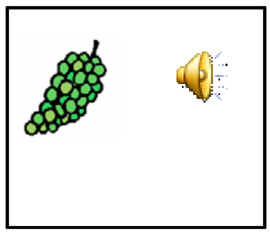
Auditory cue: It starts with /gr/. What is this?

[Slide 3]



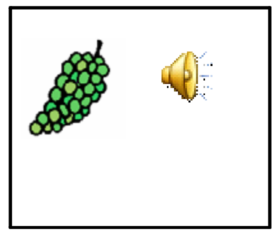
Auditory cue: It's grapes. Say grapes.

[Slide 4]



Auditory cue: Listen to me say it. Grapes (x5). Say grapes.

[Slide 5]



Auditory cue: Now say it with me. Grapes (x5). What is this?

[Slide 6]