The Effect of a Token Training Program on Auditory Comprehension in a Case of Wernicke's Aphasia

Linda Howard Burger and Robert T. Wertz
Veterans Administration Medical Center, Martinez, California

There has been no lack of interest in aphasic performance on the numerous varieties of the Token Test. Some investigators have examined the effectiveness of auditory comprehension training programs based on versions of the Token Test. West (1973) observed that mild to moderately severe aphasic patients made significant improvement on a Token Test training task; however, performance did not generalize to other language measures. Holland and Sonderman (1974) reported similar results on a Token Test treatment program for a sample of mild to moderately involved aphasic patients. However their group of severely involved patients did not improve on the treatment task. Both groups failed to generalize performance to other measures. We (Burger, Wertz, and Woods, 1983) treated a cortically deaf patient with a Token Test treatment task. Our patient improved in the treatment condition, and he displayed improved performance on other language measures administered pre- and post-treatment.

The purpose of this paper is to report the effects of a token training program (similar to the one we used with the cortically deaf patient) with an aphasic patient displaying severe auditory comprehension deficits.

CASE REPORT

J.C., a 73 year-old right-handed male, suffered a left hemisphere CVA in April 1982 during an angiogram. He received four months of individual treatment in Palo Alto Veterans Administration Medical Center immediately following his stroke. At one year postonset, J.C. was evaluated in our clinic to determine whether additional treatment would be appropriate. Performance on the Western Aphasia Battery (WAB) (Kertesz, 1982) classified J.C. as demonstrating Wernicke's aphasia. His speech was fluent and contained frequent literal and verbal paraphasias, some neologistic jargon, and strings of meaningless but grammatically connected words. Auditory comprehension deficits were severe--3.5 on the WAB ten-point scale. During conversation, it was necessary to repeat, rephrase, and demonstrate questions and comments for J.C. to comprehend.

We offered J.C. individual treatment designed to improve his auditory comprehension through the use of a systematic program based on items from the Revised Token Test (RTT) (McNeil and Prescott, 1978). In addition, he was given homework to improve writing, and he began to attend a weekly aphasia maintenance group.

Treatment Design

We began with an A-B-A single subject withdrawal design. This was eventually modified to an A-B-A-C-A design following a lack of progress during the first treatment phase. Stimuli and procedures were similar to those in the Revised Token Test. J.C. was seen twice a week in 60-minute sessions for a total of 34 sessions. Treatment began at Level I of the Revised Token Test, because J.C. experienced considerable difficulty with Level I items in a pretest. Level I of the Revised Token Test utilizes large squares and circles

in five different colors and requires a patient to indicate the correct token when instructed to, for example, "touch the blue circle." Our treatment program attempted to teach J.C. this task.

All "A," or baseline and withdrawal, phases of the program consisted of standard administrations of Level I of the Revised Token Test. No indication of the accuracy of response was given to the patient. The percent of critical elements correct was plotted in each of these sessions.

The first treatment task consisted of breaking the commands into two steps. The first step involved changing only the color element of the command. The five circles were placed on the table, and J.C. was given commands such as, "touch blue." The procedure was repeated with squares. Each trial was scored as correct or incorrect, and feedback on accuracy was provided. To monitor performance, daily criterion runs composed of the ten items in Level I of the Revised Token Test were administered at the end of each session exactly as in baseline and withdrawal.

The second step in the first treatment task involved changing only the shape element of the command. A square and circle of the same color were placed on the table and commands such as "touch the square" were given. Scoring and daily criterion runs were the same as in step one. We entered a withdrawal phase after completing step two. J.C. had met our treatment performance criterion (80 percent correct in three consecutive sessions) on each step, but he had made minimal gains on the Daily Criterion Runs.

Next, we shifted to a second treatment which began by telling J.C. which tokens were circles and which were squares and asking him to repeat the word "circle" or "square" while pointing to the respective tokens. This procedure was repeated until J.C. could say the target words correctly. Because he had attempted to repeat commands in the first treatment, we guessed that correct verbal production of the critical elements might improve performance. As in the first treatment, the program was hierarchial, however, instead of reducing the number of critical elements J.C. attended to at one time, we reduced the number of stimuli to four tokens and worked back up to ten tokens.

Step one in the second treatment consisted of placing four tokens of two different colors on the table--for example, the red circle and square and the blue circle and square. Two cards containing the printed words "square" and "circle" were placed next to the appropriate tokens as cues. J.C. was asked to "touch the red square." Feedback on accuracy was provided, and a repeat was given if a response was incorrect. If his response remained incorrect after the repeat, we showed him the correct response. When he was 80 percent correct at "11" or above on the critical elements, the cue cards were removed and the step was repeated.

Step two consisted of the same procedure using an array of six tokens. When criterion (80 percent correct on the treatment stimuli in three successive sessions) was met, two more tokens were added until an array of ten tokens was achieved. Fifty treatment trials were administered in each session, and a daily criterion run was done at the end of each session. When the 80 percent criterion was reached on ten tokens in the response matrix, we entered the second withdrawal phase.

Generalization Measures

Two generalization probes were begun in conjunction with the first withdrawal phase, just prior to beginning the second treatment. One probe consisted of ten Level I commands using five cups and plates that were the

same colors as the treatment tokens. The second probe consisted of ten Level I commands using five triangles and five ovals that differed in color from the treatment tokens. The administration and scoring of the probes were the same as in the daily criterion runs.

Language Measures

A battery of language measures composed of the Porch Index of Communicative Ability (PICA) (Porch, 1967), Western Aphasia Battery, Communicative Abilities in Daily Living (Holland, 1980), and the Revised Token Test was administered pre- and post-treatment.

RESULTS

J.C.'s performance on the treatment tasks was plotted as the percent of critical elements correct in each daily criterion run (Figure 1). Mean performance in baseline was 45% correct. Mean performance in 11 Treatment 1 sessions was 58% correct. Mean performance during withdrawal was 47%. The most appropriate adjective that describes J.C.'s performance is "variable". Inconsistency continued during Treatment 2. Mean performance across sessions was 70% correct, but this ranged from 50% to 90%. During withdrawal, mean performance was 78% correct, peaking at 95% and dropping to 55%.

Figure 2 shows comparison of performance on the treatment task and performance on the two generalization probes, begun during the first withdrawal phase. The behavior that generalized was J.C.'s inconsistency. The sawtooth pattern seen on the treatment stimuli is present in performance on both generalization probes.

Comparison of pre- and post-treatment performance on the complete Revised Token Test (Table 1) indicated a one percentile increase in J.C.'s overall score. A bit of improvement occurred on Subtests I, II, and III, the area where we concentrated treatment, but performance on the other subtests remained unchanged except for a little slippage on Subtests VII and VIII.

Table 1. Mean scores and percentiles for pre- and post-treatment administrations of the Revised Token Test.

LEVEL/SUBTEST	PRE-TREATMENT		POST-TREATMENT		DIFFERENCE	
	$\bar{X}$	%ile	$\bar{X}$	%ile	$\bar{X}$	%ile
I	12.20	(8)	12.56	(13)	+0.36	(+5)
II	11.75	(16)	11.95	(20)	+0.20	(+4)
III	8.72	(1)	9.63	(5)	+0.91	(+4)
IV	8.27	(1)	8.64	(1)	+0.37	(0)
V	5.00	(1)	7.37	(1)	+2.37	(+2)
VI	7.38	(4)	7.46	(4)	+0.08	(0)
VII	8.33	(5)	7.53	(2)	-0.80	(-3)
VIII	7.63	(7)	7.41	(6)	-0.22	(-1)
IX	5.00	(1)	6.75	(1)	+1.75	(0)
X	4.00	(1)	6.60	(1)	+2.20	(0)
Overall	7.83	(1)	8.59	(2)	+0.76	(+1)

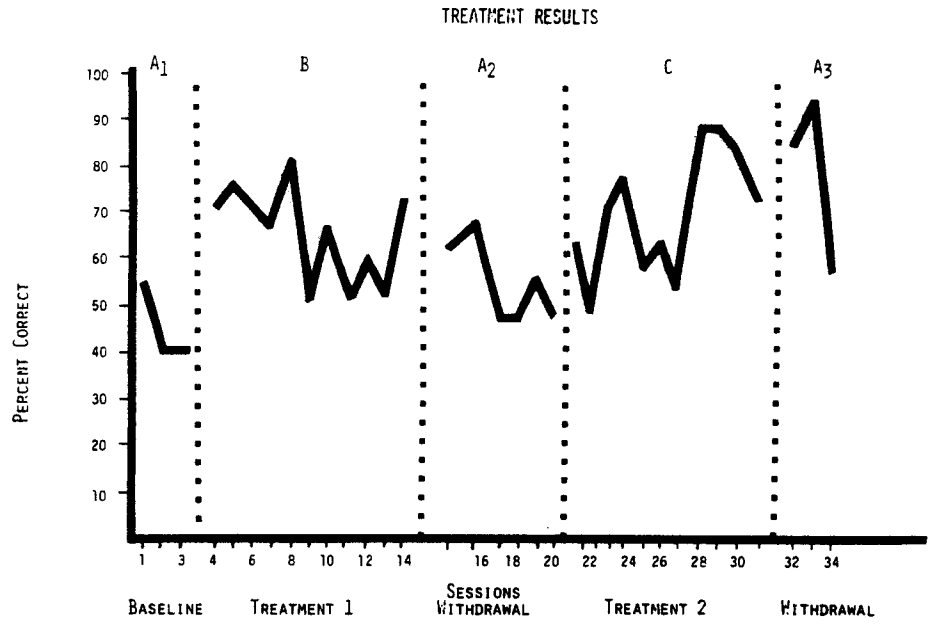

Figure 1. Baseline, treatment, and withdrawal performance on the Revised Token Test treatment task.

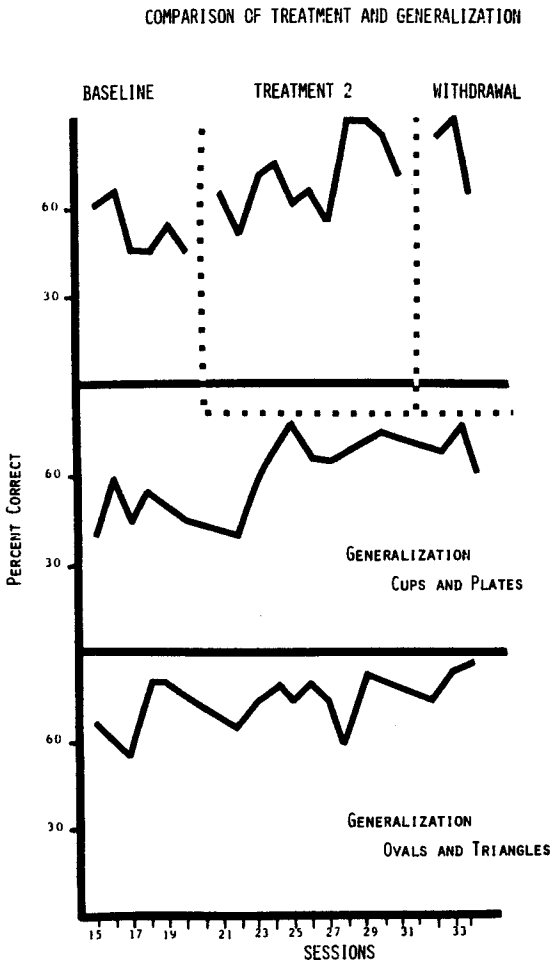

Figure 2. Performance on the treatment task and the two generalization probes.

Post-treatment performance on the Western Aphasia Battery (Table 2) indicated an improved Aphasia Quotient, improved Spontaneous Speech, and improved Auditory Comprehension. Change in the latter resulted from better post-treatment performance on yes/no questions and sequential commands. There was minimal change in auditory word recognition. Post-treatment PICA performance was up in all modalities, and the CADL total score had improved.

Table 2. Pre- and post-treatment comparisons for the Western Aphasia Battery (WAB), Porch Index of Communicative Ability (PICA), and Communicative Abilities in Daily Living (CADL).

MEASURE	PRE	POST
WAB		
Aphasia Quotient	46.80	54.60
Spontaneous Speech	14.00	16.00
Repetition	1.40	1.00
Comprehension	3.50	5.80
- yes/no questions	27 (raw scores)	42 (raw scores)
- auditory word recog.	37 (raw scores)	41 (raw scores)
- sequential commands	6 (raw scores)	34 (raw scores)
PICA		
Overall	30 %ile	41 %ile
Gestural	59	75
Verbal	44	51
Graphic	10	33
Auditory	22	41
CADL	40%	48%

DISCUSSION

We knew where we were going, and we know the path we plodded, but we do not know how or why we arrived where we did. Our treatment task was designed to improve performance on Level I of the Revised Token Test. J.C.'s erratic performance does not testify to the terrificness of our treatment. Nevertheless, he did display observable gains on the general language measures. Thus, we must attempt to explain the clinical significance of improvement in a treatment failure. Somewhere on the way from our design to our data, we defined entropy.

Perusal of the previous Token Test treatment studies provides little relief. Our cortically deaf patient (Burger et al., 1983) improved on a Token Test treatment task, and he improved on more general language measures. West's (1973) mild to moderate patients improved on her Token Test treatment task, but they did not display significant improvement on portions of the Minnesota Test for Differential Diagnosis of Aphasia (MTDDA) (Schuell, 1965). Holland and Sonderman's (1974) mild to moderate patients made gains on their

Token Test treatment task, but their moderate to severe patients did not. Neither group displayed improved performance post-treatment on selected MTDDA subtests. Again, our patient did not improve on the Token Test treatment task, but he did improve on a variety of general language measures. While the turgidities of our design differed from our previous treatment of the cortically deaf patient and from the designs employed by West and by Holland and Sonderman, the differences in designs are not sufficient to explain the differences in results.

It is true that the Token Test treatment was not the only thing happening in J.C.'s life. We provided him with homework designed to improve written skills. This may explain the 23 percentile increase in his post-treatment PICA Graphic performance. Further, he participated in a weekly maintenance group. However, our previous analysis of the influence of that group on changing PICA performance (Wertz, Collins, Kitselman, and Deal, 1983) indicated that it did not. Finally, it is not likely that spontaneous recovery was waiting to work its will on this patient at 12 months postonset.

Perhaps the Token Training treatment was facilitating and influenced performance on the WAB, PICA, and CADL but not the more difficult, nonredundant Token Test stimuli. But, as Rosenbek (Wertz, LaPointe, and Rosenbek, 1984) have observed, "how thin the air is around the word 'perhaps'." And, I would add, around the word, "facilitating."

But, someone might say, your patient was variable in his performance, and you may have administered your post-treatment language measures when he was at a peak and not in a trough. Perhaps, but the measures were administered on different days, and J.C.'s improved performance on these measures in subsequent re-evaluations post-treatment has persisted. Further, although his aphasia remains dense, his family reports that he participates more in conversations and understands them better following the Token Test treatment. Good for them.

Our results, and our failure to explain them adequately, do not end in apostasy. We continue to utilize single subject designs, and we may try another patient on a Token Test training task. However, next time, we may employ Brilliant's (1979) research maxim, "To be sure of hitting the target, shoot first and, whatever you hit, call it the target" (p. 132).

REFERENCES

Brilliant, A., I May Not Be Totally Perfect, But Parts of Me Are Excellent. Santa Barbara: Woodbridge Press, 1979.

Burger, L.H., Wertz, R.T., and Woods, D., A response to treatment in a case of cortical deafness. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1983. Minneapolis, MN: BRK Publishers, 127-136, 1983.

Holland, A.L., Communicative Abilities in Daily Living. Baltimore: University Park Press, 1980.

Holland, A.L. and Sonderman, J.C., Effects of a program based on the Token Test for teaching comprehension skills to aphasics. Journal of Speech and Hearing Research, 17, 589-598, 1974.

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

McNeil, M.R. and Prescott, T.E., Revised Token Test. Baltimore: University Park Press, 1978.

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

Schuell, H., The Minnesota Test for Differential Diagnosis of Aphasia. Minneapolis, MN: University of Minnesota Press, 1965.
Wertz, R.T., Collins, M.J., Kitselman, K.P., and Deal, J.L., Language after aphasia therapy: Maintenance of treatment gain. Paper presented to the American Speech-Language-Hearing Association, Cincinnati, Ohio, November, 1983.
Wertz, R.T., LaPointe, L.L., and Rosenbek, J.C., Apraxia of Speech in Adults: The Disorder and its Management. New York: Grune & Stratton, 1984.
West, J.A., Auditory comprehension in aphasic adults: Improvement through training. Archives of Physical Medicine and Rehabilitation, 54, 78-86, 1973.

DISCUSSION

Q: Did you consider extending your treatment phases to see if some of the variability would subside or level out?
A: We did. We talked it over with the patient, and he decided that 34 administrations of Part I of the Revised Token Test were about enough. He was ready to stop.

Q: Did you expect his behavior to reverse when you withdrew the treatment?
A: I did not; perhaps a little sag but not a sharp drop. Some insist that one of the tests of a treatment's efficacy is a reversal during withdrawal. But we like patients to remain fixed.

C: I see overlap between treatment and withdrawal phases. There is a progression from baseline to better performance during the first treatment, some maintenance during the first withdrawal, and then an upward trend during the second treatment. I guess I don't think your patient's performance was as variable and meaningless as you do.

C: Some people advocate using an equal number of baseline and treatment sessions. If you had more baseline sessions, you may have seen some reduction in the patient's variability. Also, as I remember, you had more than one task in your treatment. This may have induced some of the variability. Finally, if you had held out some of your baseline items and used these as probes across treatment days, as an extended baseline, you may have obtained an explanation of some of the variability.

C: The traditional view is that when you intervene with a treatment, it should have an immediate and dramatic effect. Otherwise, the treatment does not work. I wonder if this is true with brain damaged patients. Perhaps we need to extend our treatment over longer periods and not expect an immediate effect.

Q: Had this patient been in therapy before you began the treatment reported here?
A: He received four months of treatment immediately postonset, but he had been sitting at home for eight months prior to the treatment trial we reported.

C: I don't think I am as pessimistic as you are about your data. Even though performance is variable, there is improvement during treatment. If you used LaPointe's trend lines, I think you would see a positive treatment effect.

C: I just want to be another member of the support group and say you not only have some significant trend lines but also improvement in auditory comprehension on your general language measures. Both of these, coupled with the family reporting the patient is doing better, suggests you should cheer up.

C: I'd say cheer down. Slope, magnitude, and trend are fine, but when treatment data points overlap with baseline data points, I have difficulty seeing a positive treatment effect.

C: I disagree. Sure, there is a lot of variability, and perhaps that is the point. If this had been a group study, the variability would have been averaged out. Or, if the general language measures had been administered pre- and post-treatment and we had not seen the variability on the treatment task, we would have looked at the pre-to-post improvement and said, "look at the wonderful job they have done." Because we have the complete subject data, we can say they present some nice changes, but there are some things that need exploration and explanation.