Analysis of a Battery Assessing Mild Auditory Comprehension Involvement in Aphasia

Connie A. Tompkins, Marie T. Rau,
Robert C. Marshall, Karen J. Lambrecht
Veterans Administration Medical Center, Portland, Oregon

Lee Ann C. Golper and David S. Phillips
University of Oregon Health Sciences Center, Portland, Oregon

The need to obtain more detailed information about different aspects of auditory language comprehension is one long recognized in clinical aphasiology. Aphasia clinicians often want more information than that obtained from the administration of a standard aphasia test for establishing baseline performance, planning treatment, or counseling. The recently published Revised Token Test (McNeil and Prescott, 1978) and Auditory Comprehension Test for Sentences (Shewan, 1980) allow more in-depth measurement of the auditory capabilities of aphasic patients. Another promising instrument, proposed by Berry (1976), is a 10-task auditory supplement to the Porch Index of Communicative Ability (PICA; Porch, 1967), designed to provide important clinical information about the mildly involved patient. The battery of tasks capitalizes on the PICA's homogeneous design and sensitive scoring system, but was constructed to elaborate upon the limited information obtained from auditory subtests VI and X (identifying objects given function; identifying objects given name).

This investigation was conducted to provide specific information on Berry's proposed Advanced Auditory Battery (AAB), by addressing the following questions:

1. What are the performance levels on the 10 AAB tasks of a group of aphasic adults with high level auditory comprehension?
2. Which tasks provide information on auditory comprehension ability beyond that gleaned from PICA subtests VI and X?
3. Which AAB subtests differentiate performance of aphasic subjects from normal controls?
4. What differences, if any, exist between high comprehending fluent and nonfluent aphasic patients in performance on the AAB?
5. Can the tasks on the AAB be arranged in a hierarchy of difficulty?

METHOD

Subjects. Subjects in this study were 24 aphasic veterans (20 male and 4 female) with high-level auditory comprehension, as defined by the following performance criteria on PICA subtests VI and X:

1. no errors (scores of 8 or above achieved on each item)
2. mean scores of 13.0 or better on each subtest.

Judgments of comprehension level were made from the most recent performance on the auditory subtests of the PICA, or from a screening administration of these subtests prior to testing with the AAB. Subjects' performance for these criteria ranged from 13.6 to 15.0 with a mean of 14.33 on VI, and 13.4 to 15.0 with a mean of 14.48 on X.

Subjects ranged in age from 34 to 77 years ($\bar{x}$ = 57), in duration of aphasia from 2 to 163 months ($\bar{x}$ = 34), and in severity of aphasia as determined by overall PICA percentiles from the 50th to the 96th ($\bar{x}$ = 75.8). Etiology for 20 patients was thrombotic or embolic CVA, and hemorrhagic for two; two patients had suffered traumatic head injury. Eighteen of the subjects were classified as exhibiting nonfluent aphasia, while speech output of six was characteristic of fluent aphasia.

Three non-brain-damaged veterans from the rehabilitation ward served as controls. They ranged in age from 57 to 64 years ($\bar{x}$ = 61.3). One was a laryngectomized patient and two had below-knee amputations.

Procedures. All subjects were administered the 10-task AAB in the order suggested by Berry (1976). Table 1 provides a description of the elements involved in the 10 tasks. Standard PICA layout was used, and stimuli were delivered at a normal rate without special prosodic emphasis. Instructions and scoring followed Berry's guidelines. Four videotaped administrations were randomly selected to assess interscorer reliability. Agreement was calculated at 90%.

Table 1. Description of task requirements on AAB.

Subtest	Bits of Information	Elements
A1	6	sequence, prepositional relation, verb, three nouns
A2	4	sequence, three nouns
A3	5	sequence, two verbs, two nouns
A4	3	sequence, two nouns (delay between nouns)
A5	3	sequence, two nouns
A6	3	prepositional relation, two nouns
A7	2	verb, noun
A8	1	noun (carrier phrase)
A9	1	noun (name spelled)
A10	1	noun (5-second delay)

RESULTS

In order to determine equivalence of the fluent and nonfluent groups in terms of overall severity and auditory comprehension performance on PICA subtests VI and X, t-tests were run between the two groups on those three variables. In each case, the test was not significant at the .05 level; hence, the two types of aphasic patients did not differ on any of these measures.

A 3 x 10 Repeated Measures Analysis of Variance (Winer, 1971) with repeated measures on the second factor, was run to compare the responses of the normal, fluent aphasic, and nonfluent aphasic groups on the 10 tasks of the AAB. The Groups and the Tasks factors, as well as the Groups x Tasks interaction were significant at $p \leq .01$. Overall group mean scores on the AAB were 11.55 for fluent patients, 13.03 for nonfluent patients, and 14.89 for controls. Newman-Keuls tests showed the fluent and nonfluent groups to differ significantly from normals ($p \leq .01$), but the two aphasic groups not to differ significantly from each other. Figure 1 shows group mean performance for all 24 aphasic patients and non-brain-injured controls on the 10 tasks; Figure 2 shows group mean scores for the 2 subgroups of aphasic subjects on the battery.

To further clarify both the main effect for Tasks and the Groups x Tasks interaction, simple effects were tested for the tasks within each group. Three 1 x 12 Analyses of Variance were performed to detect differences between any of the 10 AAB tasks and the two PICA auditory subtests for each group. Significant differences were obtained for the fluent group ($F=26.86$; df, 9, 216; $p < .01$) and the nonfluent group ($F=35.98$; df, 9, 216, $p < .01$). There were no significant differences between subtests for normal controls. Newman-Keuls tests comparing performance levels on these 12 measures show similar patterns of differences for the two aphasic groups. Only five of the AAB subtests were significantly different from PICA VI and X for both groups of aphasic subjects. Table 2 shows the results of the Newman-Keuls analyses for those five subtests. One other significant difference was found for the fluent group only. These patients performed significantly more poorly on subtest A9 (pointing to items whose names are spelled) than on PICA VI and X, or on subtests A8 and A10 of the AAB.

Table 2. Summary of Newman-Keuls analysis, comparing every subtest with every other subtest, for subtests significantly different from PICA VI and X. N=significantly different for nonfluent patients; F=significantly different for fluent patients.

Subtests	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	VI	X
A1	--	N,F		N,F	N,F	N,F	N,F	N,F	N,F	N,F	N,F	N,F
A2		--	N,F	N,F			N,F	N,F	N	N,F	N,F	N,F
A3			--	N,F	N,F	N,F	N,F	N,F	N,F	N,F	N,F	N,F
A5					--			N,F		N,F	N,F	N,F
A6						--	N,F	N,F	N	N,F	N,F	N,F

Since both groups of aphasic subjects were statistically equivalent in terms of overall severity and auditory comprehension performance, the 24 subjects were pooled and compared with the control group on each of the 10

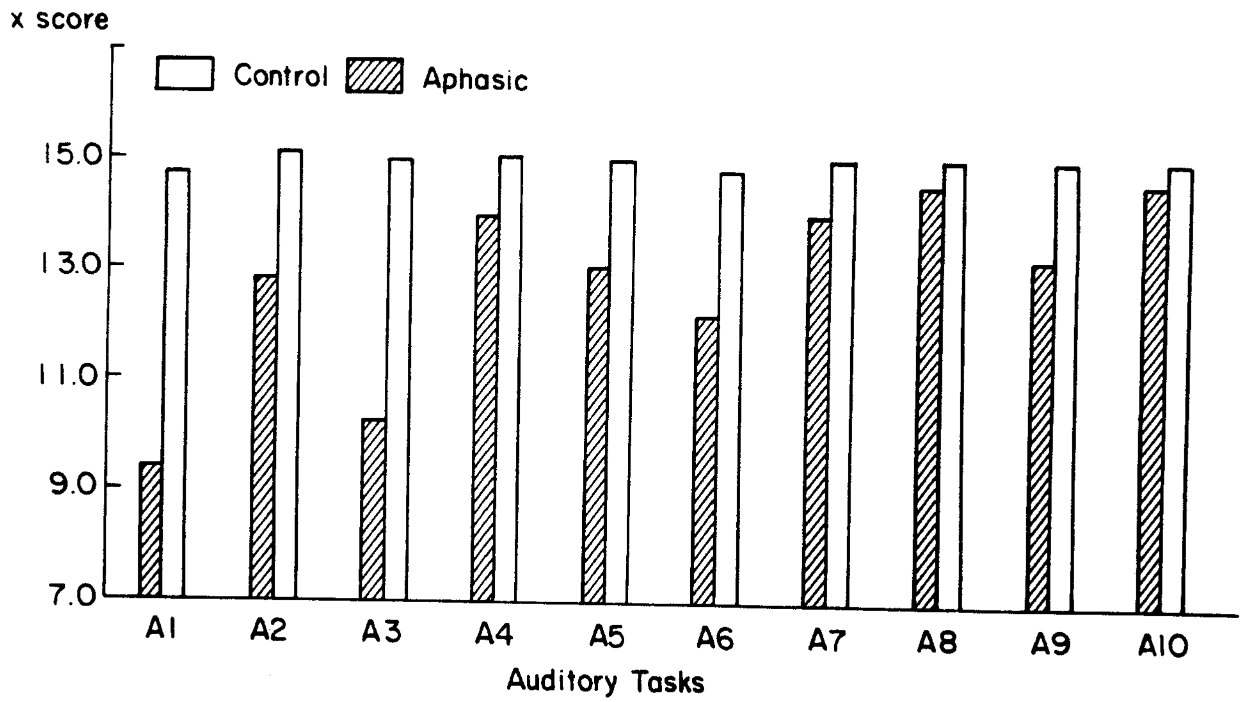

Figure 1. Group mean scores for aphasic (N=24) and control (N=3) subjects on Advanced Auditory Battery.

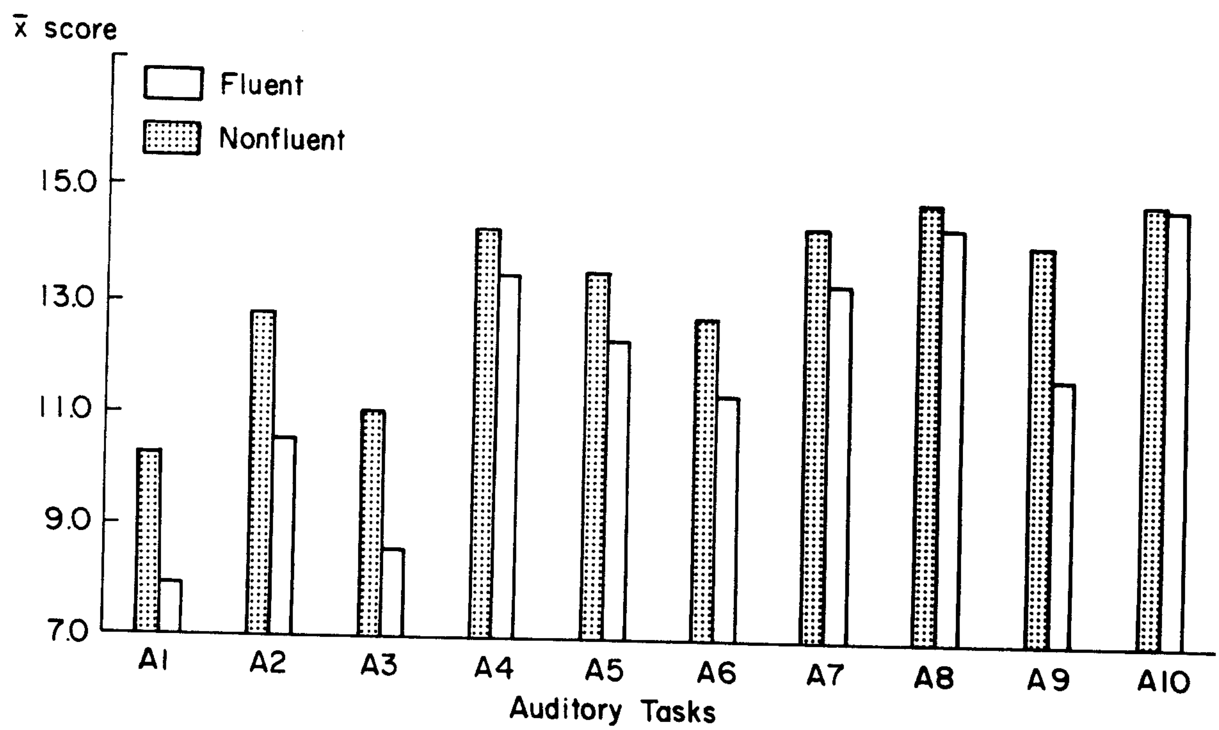

Figure 2. Group mean scores for nonfluent (N=18) and fluent (N=6) aphasic subjects on Advanced Auditory Battery.

AAB subtests. Only four tests were found to differentiate aphasic and normal performance. These are subtests A1 ($F=24.86$, df, 1, 25; $p \leq .01$), A2 ($F=7.99$, df, 1, 25; $p \leq .01$), A3 ($F=19.55$, df, 1, 25; $p \leq .01$) and A6 ($F=6.48$, df, 1, 25; $p \leq .05$). To examine the relationship between these four subtests, Pearson correlations were computed using the scores of all 24 aphasic subjects. The correlations between all subtests were consistently high, ranging from .722 to .860 (See Table 3).

Table 3. Product-moment correlation coefficients for the four discriminating AAB subtests for aphasic subjects (N=24).

Subtests	A1	A2	A3	A6
A1	--	.722	.860	.776
A2		--	.770	.795
A3			--	.785
A6				--

DISCUSSION

The results of this study indicate that some tasks from the AAB could be administered to high comprehending aphasic patients to expand upon the information obtained from the auditory subtests of the PICA. Five of the ten subtests of the AAB yielded scores significantly different from those achieved on PICA auditory tests VI and X. Only four of the five (A1, A2, A3, and A6) differentiated performance of aphasic and control subjects, however. For those four tasks, auditory retention requirements span three to six bits of information, and specification of elements in error allows determination of relative difficulties with sequential, relational or content units. While the four tasks would appear to provide a range of clinically useful information in terms of impaired comprehension processes, the high intercorrelations between the subtests cause this interpretation to be questioned. It is impossible to determine whether these high correlations indicate that the tasks are tapping the same processes, or whether the difficulty level of these tasks is consistent for subjects with good auditory comprehension.

Interpretation of the data in Table 2 and a rank ordering of the difficulty of the tasks suggests a possible hierarchy. Subtest A1 was significantly more difficult than all others, with the exception of A3. However, the rank ordering of these subtests showed that 70% of the subjects found task A1 most difficult, as compared with 33% for task A3. Similarly, tasks A2 and A6 did not differ statistically. Although A3 and A6 were statistically equivalent to A1 and A2 respectively, their inclusion would not seem redundant because of the potentially different clinical information which could be gleaned from each. In a hierarchy from most to least difficult, A1 occupies the first position, with subtest A3 second; and A2 and A6 being roughly equivalent, and in the last two positions.

Interestingly, the subgroups of fluent and nonfluent aphasic patients demonstrated similar performance on all subtests but one. Identification of objects whose names are spelled (A9) served as a discriminator between the two groups, with nonfluent subjects receiving higher scores as a group. In terms of overall performance on the battery, however, nonfluent aphasic patients did statistically no better than the fluent group.

Several possible reasons for the lack of differences between the two aphasic groups emerge. First, the scoring system, particularly for the discriminating subtests of the AAB, may disguise differences between the two groups. In general, the scores are not discriminative at lower performance levels. A score of 7 was given on subtests A1 and A3 when only one element of a command was in error, and a 6 when more than one portion of the response was incorrect. A possible solution would be the inclusion of a 50% rule similar to that used in PICA scoring, assigning a score of 4 on any occasion in which more than half of the elements of a command were in error. Use of a 4 score could result in larger gaps between mean scores for the two aphasic groups on these most difficult tasks, and might also produce significant differences between the most difficult tasks (A1 and A3) for the entire aphasic group.

A second possible problem is that the subtests were not challenging enough to discriminate between the groups. Figure 2 shows that noticeable differences between the mean scores for the two groups began to appear on subtests A1, A2, and A3. Slightly increasing the difficulty level, and/or using a revised scoring system, as suggested above, might enhance the differences between groups.

In addition, the qualitative aspects of the errors were not analyzed for the aphasic subgroups. It could be expected that the groups would differ with regard to the elements of a message posing the most comprehension difficulty. Finally, the lack of statistically significant differences may be attributable to the small number of subjects (N=6) in the fluent group.

Administration of the four tasks described in this paper may potentially offer good clinical data about mild auditory comprehension deficits in aphasia. Further research with the abbreviated battery is needed, however. Scoring guidelines require refinement and clarification, especially at the lower end of the scale. In addition, more data should be gathered with large groups of aphasic subjects for standardization and validation purposes. Applying some or all of these suggestions may allow us to determine whether this test can provide information on differential impairments of the auditory comprehension process.

REFERENCES

Berry, W., Testing auditory comprehension in aphasia: Clinical alternatives to the Token Test. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1976. Minneapolis, MN: BRK Publishers, 1976.

McNeil, M. and Prescott, T., The Revised Token Test. Springfield, IL.: Charles C. Thomas, 1978.

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

Shewan, C.M., Auditory Comprehension Test for Sentences. Chicago, IL: Biolinguistics Clinical Institutes, 1980.

Winer, B.J., Statistical Principles in Experimental Design. New York: McGraw-Hill, 1971.

ACKNOWLEDGMENT

This research was supported by the Veterans Administration and the Research and Development Committee of the Veterans Administration Medical Center, Portland, Oregon. Special thanks go to Dr. Richard Schulz for his statistical advice.

DISCUSSION

Q: Why were you interested in comparing the two groups?

A: It was a division we were able to make as we collected a large sample of patients. These fell pretty easily into groups. Auditory comprehension has been suggested as a differentiating feature of fluent and nonfluent aphasia, and we were interested to see if that would hold up at this level of comprehension ability. For these tasks, it does not.

Q: I think maybe you're a little conservative with your decision about whether the fluent people differed from the nonfluent people. I think maybe it's a case in which you get cowed by your statistics. If you look at Table 2 and do some sign tests (which don't have much power, but have intuitive appeal)--my recollection of the sign test table is if you have 10 out of 10 going in the same direction that's a very significant difference. I think if you go away from classic statistics and go to something a little more intuitively appealing and straightforward, you may find some differences that we may believe in. It looks to me as if nonfluent patients are not as severely impaired on your test as fluent patients are. That's logically appealing too.

A: An added comment, when we were looking back at the scores after noticing the problems with the scoring system, we saw that indeed 4 of the 6 fluent patients had scores on the hardest test in the "6" range, which could mean that they were getting only 1 piece of that command, or 4 out of the 6 elements. So I think perhaps a bit of scoring revision would show these differences too.

Q: Question was about composition of the fluent group, with the comment that the inclusion of a number of Wernickes patients may bring the performance levels down.

A: Our fluent patients were at fairly high severity levels, and, as you recall, were not statistically different from the nonfluent group in terms of our PICA selection criteria. We did not have a Wernickes patient in the sample. We had some who had evolved from Wernickes toward anomic, and we had some conduction aphasic patients.

Q: In that case, is making a differentiation between the aphasic subgroups an issue?

A: We were trying to collect some performance data on tasks of this level, and we thought this would be a useful division on which to have some numbers.

Q: I'm not sure your criteria would allow you to call these "mildly impaired" patients relative to verbal comprehension. If I remember the PICA norms correctly, I believe that 40 to 50% of aphasic patients score 14.3 or higher on those two subtests; which either means you can't call them mild, or you believe that 50% to 60% of aphasic patients have only a mild comprehension problem.

A: Yes, perhaps those criteria were a bit arbitrary; and what I would say rather than arguing with you over the label "mild," would be that we know that for patients with this level of comprehension ability, this is the performance we see. I think it's a good point to take into account, though, if any more application were to be done with this.

C: I think your results indicate that these tests are more sensitive than the auditory comprehension subtests of the PICA, which I believe most people think are not very good at all.

C: We wrestled with the use of the word "mild" and came up with an operational definition, and decided to include all those people for whom we got little information from VI and X.