

The Differential Diagnosis of Aphasia

Frederic L. Darley

Mayo Clinic, Rochester, Minnesota

Doctor Linebaugh has discussed six different purposes that we may have in mind in diagnostic work in aphasia. I will confine my remarks to one of these purposes, namely, sorting out patients appropriately designated as aphasic from patients with language disorders which might be confused with aphasia. What can we do and what should we learn to do better in differential diagnosis of aphasia and other communication disorders?

Scrutiny of the literature concerning various disorders that may include language dysfunction indicates that some have dealt with the issue in a gross manner, not using tests of language function but rather reporting interesting and possibly distinctive characteristics of the communication of patients they study. Thus in the literature concerning chronic schizophrenic patients we find statements that they do not use language primarily for informational purposes, that their prosody may be abnormal, that there is preoccupation with certain themes and perseveration of ideas from one context to another, that they demonstrate structurally vague or stylized constructions, etc. In discussing the differential diagnosis of dementia, Katzman and Karasu (1975) guide us in the diagnosis of Alzheimer's disease with comments like this: "There may be changes in the quality of association in the patient's speech so that words are used in a more primitive way" (p. 108). "As the illness progresses, the patient may showeccentric vocalizations" (p. 109). Patients with Wernicke-Korsakoff syndrome may "carry on a conversation with the examiner that, outside of the memory losses, is quite intact and intelligent" (p. 116). Their psychosis plateaus to a stable and chronic level whereas no such plateau is noted in Alzheimer's disease where "the patient may, in fact, lose speech and be left with perseverative vocalizations, including coprolalia" (p. 116). "Terminal mutism is considered pathognomonic for Pick's disease" (p. 117).

Fortunately, our differential diagnostic procedures have advanced beyond this state of description. Tests designed specifically for evaluation of language function have been used to delineate differences between different clinical groups. Dr. Linebaugh has shown us how some of these are really multiple-purpose tests used, if not designed, to accomplish such diverse purposes as determining the patient's level of functional communication, his prognosis, where treatment should start, and his progress in treatment, as well as localization of lesion, and, what we are talking about, differentiation of aphasia from other disorders. Tests designed to accomplish several things may not accomplish optimally any of them, including the differential function. Nevertheless, things have been learned with some of these tests that are useful to clinicians, especially those who work in medical settings where some of these determinations are attempted. Let's take a look at how some of these tests have been used.

From the title of her test, The Minnesota Test for Differential Diagnosis of Aphasia (1965), we might have thought that Hildred Schuell had in mind the development of a tool to separate aphasic patients from other language disordered patients, but in fact, the only people other than aphasic patients that she studied during the standardization process

were normal subjects. The normal subjects made few errors in comparison with the aphasic patients, and their errors were scattered randomly, yielding no particular pattern of language disability but probably indicating limitations of education. What she meant by "differential diagnosis" is the placement of aphasic patients within her subgroups on the basis of their patterns of performance and their prognosis for recovery.

Some of us (Halpern, Darley, and Brown (1973)) used an abbreviated adaptation of the Minnesota Test to see how four groups of patients would differentially perform. Patients met criteria for designation as displaying either aphasia, generalized intellectual impairment (dementia), apraxia of speech, or confused language, and all were given a 21-subtest examination. Scores on the tests yielded information about their auditory retention span, auditory comprehension, reading comprehension, naming, writing to dictation, and arithmetic. Their oral expression was scored in a manner to reflect correctness of syntax, adequacy of response, relevance of response, and fluency.

The four groups did perform differently, the aphasic patients showing the greatest impairment of language. Adequacy, auditory comprehension, arithmetic, syntax, and naming were impaired in all the disorders and did not clearly differentiate the groups. The aphasic group was particularly characterized by impairment of auditory retention and fluency. The demented group showed impaired reading comprehension and to a lesser extent impaired auditory retention, but relevance was not subnormal. The apraxia of speech group was particularly marked by a lack of fluency. In the confused language group, there was decided impairment of reading comprehension and writing to dictation, and the confused patients were most especially differentiated from the others by the frequent lack of relevance of their bizarre responses.

DiSimoni, Darley, and Aronson (1977) administered a somewhat expanded examination of this kind to another group, 27 patients with chronic schizophrenia. Their performance was quite unlike that of aphasic patients and could be differentiated from performances of the other groups as well. The schizophrenic subjects were most deviant with regard to the relevance of their responses and in reading comprehension, while naming, syntax, and the adequacy of their responses were essentially normal. Relating our findings to the duration of their problems, we concluded that their performance deteriorates with time, probably first in the direction of the pattern shown by confused patients and ultimately toward the pattern shown by subjects with dementia.

The Porch Index of Communicative Ability has been used to accomplish some of these distinctions. In the revised Volume 2 of the PICA manual, Porch (1971) presents a number of profiles which suggest language difficulty different from what he calls "aphasia without complications." He illustrates how in its purest form the pattern of apraxia of speech is characterized by very low scores on subtests I, IV, IX, and XII, the verbal subtests, with other aspects of language processing relatively intact. Similarly, patients with severe dysarthria present profiles that closely resemble those of apraxic patients, with the verbal scores disproportionately low, but with these verbal subtest scores rarely dropping below the 4 level, since the patients, though unintelligible, can generate a variety of speech sounds.

The patient who could not read or write before the onset of his acquired aphasia can also be distinguished by his PICA profiles. He rejects the reading and writing tasks, so scores on subtests V, VII, A, B,

C, and D are at the 5 level. If the patient is severely aphasic when he is tested early, this disproportion may not be evident, but as he improves, the reading and writing subtest scores remain disproportionately low in comparison to the rest of the profile.

Porch also describes the profile of the patient with bilateral involvement, who may be referred initially as a case with apparent unilateral involvement and aphasia. The distinguishing feature of this profile is a reversal of the relationship between auditory and visual task scores usually seen in aphasic patients. The bilaterally damaged patients find visual tasks (VIII and XI) more difficult than the auditory tasks (VI and X), whereas the aphasic patient will have more difficulty with the auditory tasks than with the visual matching tasks. Also the bilaterally damaged patient does less well gesturally and graphically than he does verbally.

At the 1978 Clinical Aphasiology Conference, Watson and Records (1978) reported on use of the PICA in differentiation of senile dementia from aphasia. They contrasted the performance of a group of left cerebrovascular accident patients with that of a group of patients with senile dementia or Alzheimer's disease. On all subtests the dementia group performed better than the CVA patients. On graphic tasks the dementia patients did worse on subtests E and F (visual stimulation only) than they did on subtests C and D (auditorily stimulated), while the left CVA patients did better on the visual than the auditory. Also noteworthy were the dementia group's poor performance on subtest F, copying geometric forms, and a decline on subtest V, primarily a test of reading ability.

Porec and Porch (1977) have also shown us how patients might look on the PICA if they were deliberately feigning aphasia. Among the differentiating features are the use of infantile responses (child-like articulation patterns and semantically infantile words and grammatical forms); inconsistency of aberrant behavior across subtests or within subtests; overdoing of a symptom that seemed to be truly aphasic-like, such as overuse of letter reversals; too much phonetic sophistication; yielding to time pressure and giving some kind of response rather than remaining unable to respond at the end of allotted time; and substituting more difficult words when simulating trouble remembering frequently occurring words. They reported that it is not so much specific behaviors that separate aphasic from non-aphasic patients but rather the pattern of scores across the whole battery of subtests that is "really discriminating of non-organicity."

Keenan and Brassell (1972) have reported how a writing test can differentiate between aphasic patients and people who are poorly educated. They compared 15 mild aphasic patients with 11 prisoners with education below the fourth grade in writing five sentences to dictation. The patterns of error were different in the two groups. Nearly three-fourths of the responses produced by the poorly educated group were either essentially correct or characterized by simple misspellings. But over three-fourths of the responses of the aphasic group had more serious errors which altered or lost the meaning of the sentences. Only the aphasic subjects made errors of omission of content words. A surprising percentage (27%) of the responses of the poorly educated were nonsense. There was some overlap of the two groups, both groups producing about the same percentage of correct sentences.

The Token Test is so discriminating as a detector of mild input problems that we are sometimes tempted to think of it as a differential test for aphasia. Regrettably it is not. There is some evidence that age

and amount of education influence Token Test performance. Boller and Vignolo (1966) have reported that Token Test scores are not influenced by intellectual deficit; they found no significant correlation between Token Test performance and performance on Raven's Progressive Matrices. But some of us have found incontrovertibly that patients who demonstrate cognitive slippage do poorly on at least parts IV and V of the Token Test. Results of administration of the Token Test to right hemisphere damaged patients are contradictory. Aten, Wertz, and Collins (1972) found that left hemisphere patients scored significantly lower on all parts of the Token Test than right hemisphere patients and normal subjects, with no difference between the right hemisphere and the normal subjects. But Swisher and Sarno (1969) found that right hemisphere patients did significantly more poorly overall and on parts IV and V than the normal group, a finding also reported by McClellan, Wertz, and Collins (1973).

It is tempting to think of the Word Fluency Measure suggested by Borkowski, Benton, and Spreen (1967) as a test for diagnosis of aphasia, but it isn't. Wertz and his associates in the completed VA Cooperative Study did find that the Word Fluency Measure was an index of improvement in recovery from aphasia; their patients in both individual and group treatment displayed significant linear improvement in word fluency scores during the first year post onset. But the test was designed to detect brain injury, and apparently that is what it does. Many patients with problems other than language-specific impairment produce few words on the Word Fluency Measure.

We can see from the foregoing that some tests may provide helpful clues in differentiating aphasia from at least some other kinds of problems. It is also quite evident that a test designed to tell us a lot about a patient's level of functional language will require him to engage in tasks that tell us little of differential diagnostic importance. Most tests are loaded with nondifferentiating material. Even if we learn to detect which subtests are most telling, it seems that our use of this information depends on our clinical judgment and the amount of clinical experience we have had. The results often turn out to be less than completely reliable.

A somewhat different approach is suggested by a report by Chédru and Geschwind (1978) in their attempt to describe what happens to the higher cortical functions in acute confusional states. They studied 24 patients who presented as a main characteristic "a reduction and/or an erratic shifting of attention." Other frequent symptoms not essential for the diagnosis included slowness and sluggishness of response, disorientation for time and space, changes in mood, hallucinations, and bilateral slow activity in the EEG. Their battery of tests including the following: an evaluation of mood, psychomotor activity, concern for illness, and presence of confabulation and hallucinations; ten questions about orientation in time and space; two tests of attention (a digit span test and a test of auditory-motor attention where letters were presented orally in random order and the subject had to indicate every time he heard the letter A); language tests including the following: spontaneous language (talking about his job and what he would do in changing a flat tire), naming 15 items, a word-list test (listing the names of animals as rapidly as possible), oral spelling of four words, comprehension of three orders similar to commands in the Token Test and three grammatically complex questions, repetition of four sentences, reading of four sentences, writing to dictation, and copying a sentence; an evaluation of right-left orientation; an evaluation

of finger recognition; three arithmetic problems, copying a cube and two designs from the Bender-Gestalt test; a study of praxis (bucco-facial, symbolic movements like a salute, movements representing the use of an object, and prolonged sequential activity with object manipulation); Head's hand-eye-ear test (e.g., touch your right ear with your left hand); visual recognition tests (describing the pictures used in the naming test and recognizing hidden figures); a memory test of general information and recall after ten minutes of three words; and two tests of abstract thought (six similarities from the Wechsler-Bellevue test and five proverb interpretations). They presented quantitative data concerning performance on the various tests, reporting, for example, that the majority of subjects had a great deal of trouble with six tests in particular: temporo-spatial orientation, the two tests of attention, writing, the word-list test, and Head's test. On nine other tests fewer of them demonstrated less severe difficulty, and on the rest of the tests the patients demonstrated essentially no deficit. These investigators also present qualitative data about how the patients went about performing and what kinds of aberrant behaviors were noted. They readministered all the tests after the patients were pretty well recovered and reported that most of the disorders disappeared completely after the acute confusional state had cleared up.

Chédru and Geschwind did not administer this battery of tests to any group other than patients demonstrating acute confusional state. If only they had! We can see that we might learn quite a bit about differential diagnosis of various disorders involving language dysfunction by using a wide-ranging battery of this kind, more wide-ranging than tests designed specifically for testing aphasia. This is exactly the point made by Orgass and Poeck (1969) in their presentation in Cortex ten years ago entitled "Assessment of Aphasia by Psychometric Methods." They agree that it would be nice to have a perfect test, one that will tell us all about the clinical characteristics of our patients and also accurately differentiate aphasic from nonaphasic patients. They call the profile analysis of aphasia tests a "clinical" rather than a strictly psychometric method. They suggest that a better procedure is to use a statistical technique such as discriminant analysis in order to determine what variables predict membership in a given diagnostic group and the respective weights of these variables. We could end up, then, not administering a large battery of subtests but fewer subtests of known discriminating power.

Orgass and Poeck (1969) made an attempt to do this on a small scale. They studied 18 left hemisphere damaged patients, six right hemisphere damaged patients, and 26 bilaterally damaged patients to see how well some selection of tests could identify the 15 aphasic patients. They put together three groups of tests: first, some tests of intelligence and memory, including the German version of the Wechsler Adult Intelligence Scale, two tests of verbal learning of paired associates, three tests of verbal retention (series of numbers, series of monosyllabic first names, and groups of words similar to items in the Token Test such as "The great blue angle and three little dots"), and Benton's Visual Retention Test, multiple choice form. Second, a group of tests of abilities frequently impaired in aphasia, including the copying of a 60-syllable paragraph and ten single rare Latin words, three subtests involving writing to dictation (a short paragraph, those same ten Latin words, and 30 2-4 digit numbers), some arithmetic tasks (12 addition and subtraction items, 12 multiplication and division items, and the Wechsler Arithmetic Reasoning subtest. Third,

they included a group of tests originally conceived to differentiate aphasic from nonaphasic patients, including Head's hand-eye-ear test (24 items in various crossed and uncrossed body tasks, 12 with the eyes open and 12 with the eyes closed) and the Token Test.

Having gotten their test results, they correlated them and grouped those scores that were highly correlated, ending up with 11 variables. They ranked the 50 subjects on these variables, then summed the ranks, then developed a new set of ranks from one to fifty on the basis of these sums. They divided the total group of 50 into the poorest scores, medium scores, and the highest scores. The Token Test and the WAIS Verbal IQ best differentiated the three levels of subjects. They compared the scores of aphasic and nonaphasic subjects using the Mann-Whitney U test, and got z values and levels of significance. There were no significant differences between the two groups except on the copying task. Since the aphasic patients did more poorly than the other patients on even nonverbal tests, they wondered whether they had a lower general intelligence level. So they correlated all the scores with Full-scale IQ, Verbal IQ, and Performance IQ. They found all tests to be related to a certain degree to general intelligence. So they ended up selecting 15 aphasic and 15 nonaphasic subjects, matched for age, educational level, and lateralization of brain lesion. All the test differences now were significant except for the copying task and the Benton Visual Retention Test. They found that the most significant differences between the two groups were on the Token Test, arithmetic, Verbal IQ, Head's hand-eye-ear test, and writing to dictation. They then took these variables and combined them step-wise: Token Test plus arithmetic for example; then Token Test plus arithmetic plus Verbal IQ; then Token Test plus arithmetic plus Verbal IQ, plus Head's test, etc.

They found that the maximum differentiation between the two groups was reached when they used four variables: Token Test, arithmetic, Verbal IQ, and the Head Test. Adding more variables did not improve the discriminating power of the battery; in fact, it reduced it. They found that the total score in this four-part battery corresponded closely to clinical diagnosis. Only two out of 15 aphasic patients would have been erroneously classified as nonaphasic. The total score ended up being better than any of its single components in discriminating between the groups.

Well, this is just what we haven't really done: to develop a discriminating, economical test which can tell us with reliability into what diagnostic group a patient should fall. Orgass and Poeck's effort was only a start; it did not include all possible variables; for example, they had no expressive verbal behavior built into their battery.

I would hope that as we look to the future in aphasia testing, this is one direction we might go, not just to keep on using the tests we have which seem to be particularly useful in telling us about a patient's performance, his prognosis, and what degree of progress he demonstrates on serial testing, but rather to go to the construction of differential diagnostic tests, tests trimmed down to those components that allow most precise differentiation of aphasic patients from illiterates, aphasic patients from demented patients, aphasic patients from confused patients, aphasic patients from dysarthric and anarthric patients, aphasic patients from right hemisphere damaged patients, aphasic patients from psychotic patients, and aphasic patients from malingerers (if there are any).

References

- Aten, J.L., Wertz, R.T., and Collins, M.J. The effects of interhemispheric lesions on nonverbal and verbal visual and auditory sequential behavior. Paper presented at the 10th Annual Meeting, Academy of Aphasia, Rochester, New York, October, 1972.
- Boller, F. and Vignolo, L.A. Latent sensory aphasia in hemisphere-damaged patients: An experimental study with the Token Test. Brain, 89, 815-830 (1966).
- Borkowski, J.G., Benton, A.L., and Spreen, O. Word fluency and brain damage. Neuropsychologia, 5, 135-140 (1967).
- Chédru, F. and Geschwind, N. Disorders of higher cortical functions in acute confusional states. Cortex, 8, 395-411 (1972).
- DiSimoni, F.G., Darley, F.L., and Aronson, A.E. Patterns of dysfunction in Schizophrenic patients on an aphasia test battery. Journal of Speech and Hearing Disorders, 42, 498-513 (1977).
- Halpern, H., Darley, F.L., and Brown, J.R. Differential language and neurologic characteristics in cerebral involvement. Journal of Speech and Hearing Disorders, 38, 162-173 (1973).
- Katzman, R. and Karasu, T.B. Differential diagnosis of dementia. Pp. 103-132 in Fields, W.S. (Ed.), Neurological and Sensory Disorders in the Elderly. New York: Stratton Intercontinental Medical Book Corp., 1975.
- Keenan, J.S. and Brassell, E.G. Comparison of minimally dysphasic and minimally educated subjects in a sentence writing task. Cortex, 8, 93-105 (1972).
- McClellan, M.E., Wertz, R.T., and Collins, M.J. The effects of interhemispheric lesions on central auditory behavior. Paper presented at ASHA Convention, Detroit, Michigan, October, 1973.
- Orgass, B. and Poeck, K. Assessment of aphasia by psychometric methods. Cortex, 5, 317-330 (1969).
- Porch, B.E. Porch Index of Communicative Ability. Vol. 2. Administration and Scoring. Rev. Ed. Palo Alto: Consulting Psychologists Press (1971).
- Porec, J.P. and Porch, B.E. The behavioral characteristics of "simulated" aphasia. Pp. 297-301 in Brookshire, R.H. (Ed.) Clinical Aphasiology Conference Proceedings, 1977. Minneapolis: BRK Publishers (1977).
- Schuell, H. The Minnesota Test for Differential Diagnosis of Aphasia. Minneapolis: University of Minnesota Press (1965).
- Swisher, L.P. and Sarno, M.T. Token Test scores of three matched patient groups: Left brain-damaged with aphasia; right brain-damaged without aphasia; non-brain-damaged. Cortex, 5, 264-273 (1969).
- Watson, J.M. and Records, L.E. The effectiveness of the Porch Index of Communicative Ability as a diagnostic tool in assessing specific behaviors of senile dementia. Pp. 93-105 in Brookshire, R.H. (Ed.) Clinical Aphasiology Conference Proceedings, 1978. Minneapolis: BRK Publishers (1978).