Foreign Language Comprehension of Aphasic and Normal Subjects

R.C. Marshall
C.A. Thompson
M.T. Rau

Veterans Administration Hospital
and
University of Oregon Health Sciences Center
Portland, Oregon

The auditory comprehension difficulties of aphasic adults have been a fruitful research area for aphasiologists for several years. Some investigators have compared the auditory comprehension performances of aphasic and normal individuals (Goodglass, Gleson, and Hyde, 1970; Shewan and Canter, 1971; Shewan, 1976). Others have attempted to differentiate the comprehension patterns for various diagnostic categories of aphasic individuals (Paresi and Pizzamiglio, 1970; Pizzamiglio and Appicciafuoco, 1971; Poeck, Kerschensteiner, and Hartje, 1972; Boller and Green, 1971; Heilman and Scholes, 1975). For the most part these studies have demonstrated that aphasic and normal adults and different categories of aphasic adults illustrate quantitative but not qualitative differences in comprehensional ability. Simply stated, this finding means aphasic adults make more errors on a comprehension task than normal adults, but items or structures likely to be misunderstood by aphasics also have a higher probability of being misunderstood by normals. Similarly, certain groups of aphasic adults, depending on degree of language involvement, site and size of lesion, and type of aphasia, make more errors than others; but do not show dissimilar error patterns or qualitative differences in error distribution.

While the studies reviewed have varied slightly in terms of populations sampled, tasks employed, and experimental methodologies, the results have been strikingly similar. The point of view adopted in undertaking the present study was that lack of qualitative differences in comprehension performance between aphasic and normal adults could be related to problems in designing a comprehension taks difficult enough to tax the systems of normal adults, yet simple enough to yield a representative performance from aphasic subjects. It is possible that the familiarity, redundancy, and suprasegmental features inherent in normal and aphasic adults' native languages coupled with the modifications necessary for designing a task suitable for both groups may preclude the illustration of qualitative differences in comprehension between them. We approached this problem by using a linguistic code that was equally unfamiliar to aphasic and normal subjects, a foreign language, specifically German.

The present study compared the performances of aphasic and normal adults on a simple foreign language comprehension program. Two questions were explor

1. Do aphasic and normal adults illustrate quantitative differences in comprehension of German?
2. Do aphasic and normal adults exhibit qualitative differences in comprehension of German?

MATERIALS AND METHODS

Subjects. Subjects for this study were 12 aphasic and 12 normal adults. Aphasic subjects ranged in age from 45 to 80 years (Mean = 58 years), in duration of aphasia from 3 to 71 months (Mean = 34 months), and in years of education from 8 to 16 years (Mean = 11.4 years). All had suffered a single thrombotic or embolic stroke and were considered to be nonfluent or Broca's aphasic (Goodglass and Kaplan, 1972). Estimates of aphasic involvement derived from subjects' overall percentile rankings on the Porch Index of Communicative Ability (Porch, 1967) ranged from the 71st to the 96th percentile (Mean = 81st percentile). The 12 normal subjects approximated the aphasic subjects in age (Range 33-79 years; Mean, 56.3 years) and years of education (Range 8-16 years; Mean, 11 years). None of the subjects reported evidence of hearing or visual difficulties. None had been exposed to German, the foreign language used in this study.

Task Materials. The experimental task for each subject was to listen to the first volume of a German language learning program developed by Winitz, Reeds, and Garcia (1975). This program, designed to teach German as a second language, is presented in small, easily mastered steps and requires simply that the learner (in this case the subject) look at the manual and point to what he hears as the stimulus. No verbalization is required. The complete program contains four 200-item volumes (800 items total) beginning with presentation of single nouns (e.g. "cat") and culminating with presentation of complex declarative sentences (e.g. "The doctor is buying the shirt also."). This study employed only the first 200 items of the program and spanned a range of difficulty from single nouns to simple declarative sentences of a subject-verb-object form (e.g., "The king is eating the pear.").

Procedures. After being familiarized with the task, each subject was individually administered the 200 item program. The first 100 items were administered, the subject was given a five minute rest, and the second 100 items were given. All stimuli were delivered by cassette tape recorder (Sony TC-150) using the tapes developed for the program by Winitz, Reeds, and Garcia (1975). For each item the manner of taped presentation was similar. A number in English cued the subject, and the stimulus was delivered. A two-second pause and repetition of the stimulus followed (e.g., Number one "cat" (pause) "cat"). The time between stimulus presentations was approximately seven seconds. Subjects were permitted to respond anytime after the first stimulus presentation. If the subject had not responded after movement of the tape to a point where the next item was about to be delivered, the experimenter (C.T.) stopped the recorder and waited for the subject to respond.

Classification of Program Items. One hundred forty-eight of the 200 items in the program required subjects to display some degree of comprehension ability. These items were classified on the basis of Item Types (noun or sentence), Item Lengths (1 = single noun or subject-verb sentence; 2 = two nouns or subject-verb-object sentence), and Item Choices or the number of pictures veiwed by the subject when listening to the stimulus (Two or Three). The eight combinations arising from this classification included noun, length-one, two-choice (N12); noun, length-one, three-choice (N13); noun, length-two, two-choice (N22); noun, length-two, three-choice (N23); sentence,

length-one, two-choice (S12); sentence, length-one, three-choice (S13); sentence, length-two, two-choice (S22); and sentence, length-two, three-choice (S23). Examples of these items and the number of occurrences for each item in the program are given in Table 1. Fifty-two program items were eliminated from the classification. These included 40 presentations

Table 1. Classification of Stimulus Items and Number (in parentheses) of Items for Each Classification Within the Program. Examples Reflect the Pictures Seen by Subjects. Stimuli were Delivered in German.

	Two-Choice	Three-Choice
Noun Items		
Length one	(25) pear cat	(46) doctor sun cat
Length two	(21) The doctor and the girl The king and the girl	(9) The king and the girl. The king and the doctor. The doctor and the girl.
Sentence Items		
Length one	(14) The doctor is running. The king is running.	(15) The girl is sleeping The girl is running. The king is sleeping.
Length two	(10) The doctor is drinking coffee The king is drinking coffee	(8) The girl is eating the pear. The doctor is eating the pear. The girl is eating the banana.

where the subject viewed a single picture and could not err, and 12 items where a choice between two pictures might be made on the basis of length rather than content of the stimulus. Although subjects listened to all 200 program stimuli and attention to these 52 items was probably necessary for success on other items, these items were not included in the data analysis.

RESULTS

All subjects responded to 200 items. Responses were scored as correct or incorrect. None of the aphasic or the normal subjects made any errors on the 52 items excluded from the data analysis.

Effects of Groups, Item Types, Item Lengths, and Item Choices. Data were submitted to a four factor analysis of variance which ascertained the effects

of groups (Aphasic and Normals), Item Types (noun or sentence, Item lengths (Length 1 or Length 2), and Item Choices (two or three) as well as the interactions between these variables. Percentages of errors for individual aphasic and normal subjects are given in Table 2. Aphasic subjects made

Table 2. Percentages of Error for Aphasic (N - 12) and Normal Subjects (N - 12) for German Language Learning Program.

Subject	Aphasic	Normal
1	32.43	2.00
2	15.54	8.11
3	31.08	6.76
4	32.43	12.84
5	9.46	0.00
6	18.92	10.81
7	20.95	7.43
8	43.92	.007
9	34.46	4.05
10	20.95	1.35
11	8.11	6.08
12	16.89	8.11
Group Mean	21.23	5.04

slightly more than four times as many errors as normal subjects. Differences between groups were significant ($F = 26.29$; $df = 1, 191$; $P < .01$) and highly supportive of previous research showing normal adults to be quantitatively superior in comprehension ability. Aphasic and normal subjects' mean error percentages for Item Types, Item Length, and Item Choices are shown in Table III. Both groups tended to make slightly more errors on sentence than noun items, but these differences were not significant ($F = 4.21$; $df = 1, 168$; $P > .05$). Both groups had significantly higher error percentages for three-choice than two-choice items ($F = 57.51$; $df = 1, 22$; $P < .01$). The group X choice interaction was significant ($F = 15.89$; $df = 1, 22$; $P < .01$) as aphasic subjects made slightly more than twice as many errors on three-choice as two-choice items; normal subjects, on the other hand, made nearly three times as many errors on the three-choice items. Subjects also made significantly more errors on length one than length two items ($F = 31.96$; $df = 1, 22$; $P < .01$). The group X length interaction was also significant ($F = 8.71$; $df = 1, 22$; $P < .01$) as normal subjects illustrated proportionately fewer errors than aphasic subjects for length two items.

Other Significant Interactions. Subjects' mean error percentages for the type X length and choice X length interactions are given in Tables 4 and 5 respectively. Table 4 shows that subjects made almost three times as many errors on length one sentence items as length two sentence items, but that the difference for noun items of different lengths was quite small. Table 5 shows that subjects had substantially higher error percentages for length one,

three-choice than length two, three-choice items, whereas the difference in mean error percentage for two-choice items of different lengths was quite small.

Qualitative Differences. Aphasic and normal subjects' mean error percentage for stimulus item types are rank ordered in Figure 1. Both groups made the

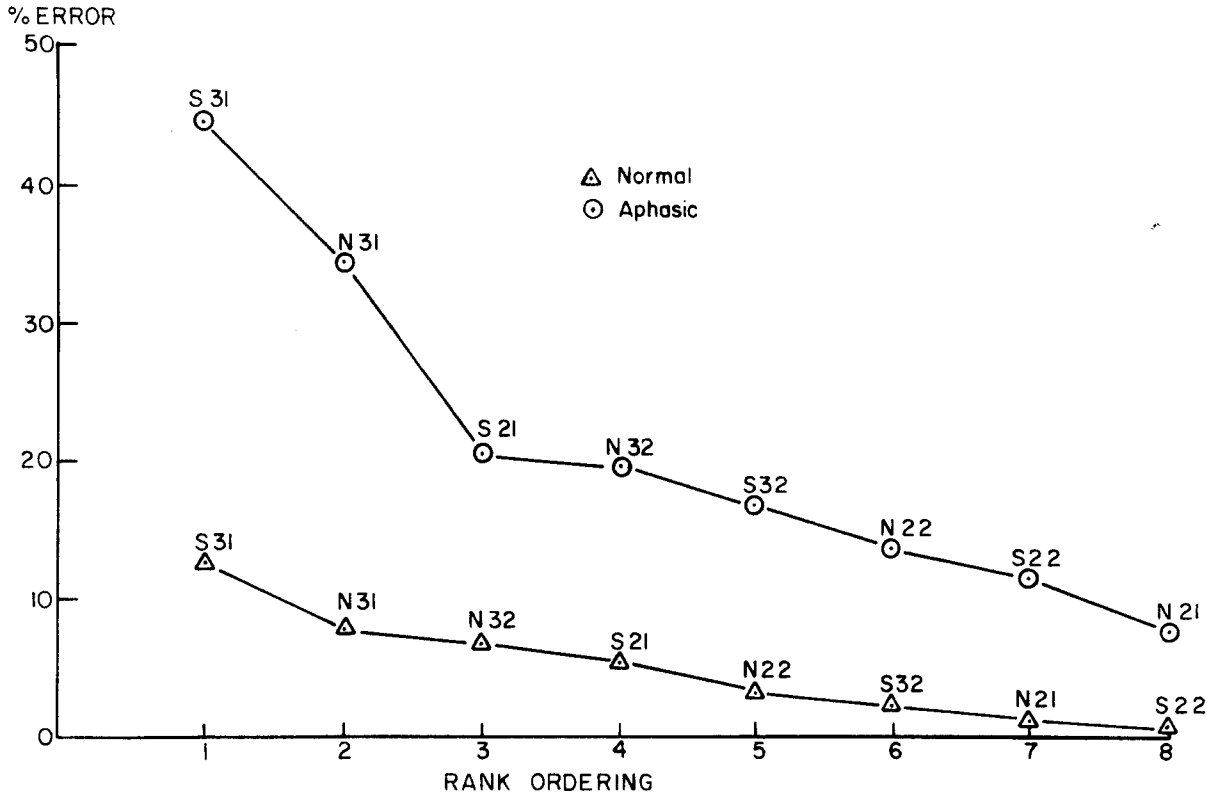

Figure 1. Error percentages for stimulus item types.

greatest percentage of errors on S31 items; both had the next highest percentages of error for N31 items. Order of difficulty (in terms of mean percentage of error) is similar in both groups for remaining items but reversed by one place for the normal group. The S21 items, third in difficulty for aphasics, ranked fourth for normals; the N32 items, fourth in difficulty for aphasic subjects, ranked third for the normals. Similar patterns prevail for items ranked fifth and sixth and seventh and eighth respectively. The pattern of errors is remarkably similar for the groups. S31 items stand out as the most difficult stimulus; there is a sharp drop (less so for normals) for N31 items with normal subjects showing a gradual decrease in mean percentage of errors across remaining items. Aphasic subjects show another sharp drop in error percentage for S22 items before following a pattern similar to normals. The results suggest that while aphasic subjects make substantially more errors than normal subjects, for all stimulus items, the same types of stimulus items tend to be difficult or easy for both groups. In short, they do not appear to differ qualitatively in ability to comprehend simple German constructions.

Other Findings. All subjects responded to the 200 items presented. The manner in which aphasic and normal subjects completed these responses was quite different, however. Aphasic subjects almost always waited for the second stimulus repetition before responding. Normal subjects usually responded after the first presentation and illustrated signs of impatience and boredom while waiting for the next item. Processing delays were defined as any undue hesitation in responding after repetition of the stimulus. One hundred seven of these delays were tabulated for the aphasic group; only seven delays were recorded for the normal group (primarily because they responded after initial presentation and did not wait for the repeat.)

Groups also tended to differ in awareness of and ability to self-correct error responses. Thirty-three self correction behaviors were recorded for the aphasic group; 45 were recorded for the normal group. While these figures represent small absolute differences, they must be viewed in the context of the total number of errors for each group. Normal subjects made 101 actual errors and 45 self corrections; aphasic subjects made 422 actual errors and 33 self corrections. By summing the actual errors and self-corrections (total errors) normal subjects are seen to self correct 30% of the time, aphasic subjects slightly less than 8% of the time.

Subjects also differed in their ability to employ associations to assign meaning to German vocabulary items. Normal subjects tended to relate German words to familiar wounding English words. For example subjects would state that the German words for "house" (haus), "hat" (hoot), and "milk" (milch) resembled their English equivalents. Some of the associations of the normals were unique such as "gabel" (fork) with gobble the food and "baum" (tree) with bough of a tree. This behavior was totally lacking for aphasic subjects, except for one subject whose performance (9.46% errors) resembled that of the normal group.

DISCUSSION

This study compared the ability of aphasic and normal subjects to comprehend German, a language equally unfamiliar to subjects in each group. We questioned whether aphasic and normal adults would differ quantitatively in comprehension of simple German structures. Results indicate that the answer is yes. We also questioned whether or not these groups would illustrate dissimilar error patterns for German structures heard for the first time. The answer to this question appears to be no. Findings are thus supportive of previously cited studies using subjects' native languages and showing that aphasic and normal subjects differed quantitatively but not qualitatively in comprehension ability.

It was anticipated that subjects might perform more poorly on longer stimulus items. This was not the case. Both groups had highest error percentages for S31 and N31 items. The tendency to make more errors on shorter items may be due to the fact that these items require the subject to discriminate between two or three single words. Longer items provide more information to guide the subject in making his selection. Hence the difference between length one and length two stimuli seems related to the factor of redundancy. Interestingly, both aphasic and normal subjects make use of this extra information, even though it is presented in a language unfamiliar to them. The groups X length interaction was significant and it would appear (Table 3) that normal subjects make greater use of redundancy as a cue by

proportionately decreasing their errors for length two items as compared with length one items.

Table 3. Aphasic and Normal Group Mean Error Percentages for Item Types, Item Lengths, and Item Choices.

	Aphasic	Normal
Item Types		
Noun	18.81	4.81
Sentence	20.80	5.26
Item Lengths		
Length one	27.03	6.86
Length two	15.42	3.21
Item Choices		
Two	13.66	2.67
Three	28.80	7.40

As expected, subjects made more errors on simple German sentences than nouns. Differences in comprehension for these structures were not significant, however. The type X length interaction was significant and it would appear that length affects subjects' ability to comprehend noun and sentence structures differentially. Table 4 shows that subjects differed only slightly

Table 4. Group Mean Error Percentages for Type X Length Interaction

	Length 1	Length 2
Noun	12.80	10.81
Sentence	21.09	7.82

(2.1% in performance) for noun stimuli of different lengths, but showed a much higher percentage of error (14%) for sentence length one items opposed to length two items. This may be due to the fact that subjects have more difficulty assigning meaning to German verbs than German nouns.

Also, as anticipated, subjects made significantly more errors on three-choice items than two-choice items. This is partially explained on the

basis of the probability of error being .66 for the former and .50 for the latter. Normal subjects (Table 3) made very few errors (2.67%) on two-choice items. Their nearly three-fold increase in errors for three-choice items is responsible for a significant group X choice interaction. While three-choice items were markedly more difficult for both groups, the magnitude of the difference was tempered by a significant choice X length interaction. Subjects made substantially more errors on three-choice, length-one items than three-choice, length-two items; differences for two-choice items of different lengths were small (Table 5). These tendencies are felt to be related to a combination of decreased redundancy and increased probability of error for S31 and L31 items. This is substantiated by the fact that subjects made their highest number of errors on these items.

Table 5. Groups' Mean Error Percentages for Choice x Length Interactions.

	Length 1	Length 2
Two choice	8.95	7.39
Three choice	24.95	11.24

Some practical clinical implications arise from this study. Findings indicate that mildly impaired nonfluent aphasic subjects are markedly impaired compared with normal subjects in their ability to understand an unfamiliar language. The magnitude of difference between these groups was somewhat surprising in light of the moderate-to-mild involvement of the aphasic subjects and the fact that the material presented in the program was elementary, often requiring only a differentiation of two items. One interpretation for such a finding might be that our traditional measures of aphasic subjects' comprehension offer too many redundancies and familiar cues to these patients to allow differentiating them from normal subjects. In this vein, the foreign language task may have provided a challenge similar to the Token Test (DeRenzi and Vignolo, 1962) and ferreted out comprehension problems not evidenced in clinical testing. Another interpretation might be that exposure to a foreign language constitutes a new learning task that is not easily mastered by aphasic individuals. This does not seem to be the case for some aphasic subjects, however, as two of the subjects performed as well as normals and the performance of two others approximated that of the normal group. In addition, the superior responses of subjects to longer more redundant structures and the tendency of aphasic subjects to take more time and require more stimulations before responding suggest these factors should be incorporated into tasks designed to provide auditory stimulation in therapy settings.

REFERENCES

Boller, F. and Green, E. Comprehension in severe aphasics, Cortex, 10, 382-394 (1971).

Derenzi, E. and Vignolo, L.A. The Token Test: A sensitive test to detect receptive disturbances in aphasics, Brain, 85, 655-678 (1962).
Goodglass, H., Gleason, J. and Hyde, M. Some dimensions of auditory comprehension performance in aphasia, J. Speech Hearing Res., 13, 595-606 (1970).
---- and Kaplan E. The Assessment of Aphasia and Related Disorders, Lea and Febiger, Philadelphia, (1972).
Heilman, K. and Scholes, R. The nature of comprehension errors in Broca's conduction, and Wernicke's aphasics, Cortex, 12, 258-265 (1970).
Paresi, D. and Pizzamiglio, L. Syntactic comprehension in aphasia, Cortex, 6, 204-215 (1970).
Pizzamiglio, L. and Appicciafuoco, A. Semantic comprehension in aphasia, J. Comm. Dis., 3, 280-288 (1971).
Poeck, K., Kerschensteiner, M. and Hartue, W. A quantitative study on language understanding in fluent and nonfluent aphasia, Cortex, 8, 299-303 (1972).
Porch, B.E. Porch Index of Communicative Ability, Consulting Psychologists, Palo Alto, (1967).
Shewan, C.M. Error patterns in auditory comprehension of adult aphasics, Cortex, 12, 325-336 (1976).
---- and Canter, G. Effects of vocabulary, syntax, and sentence length on auditory comprehension in aphasic patients, Cortex, 7, 209-226 (1971).
Winer, B.J. Statistical Principles in Experimental Design, McGraw-Hill, New York (1971).
Winitz, H., Reeds, J. and Garcia, P. Natural Language Learning, General Linguistics, Kansas City (1975).

DISCUSSION

Q. Could you explain briefly what you mean by redundancy in the short vs. long stimulus?
A. The term redundancy was used to mean "more information" on which the subjects could base a decision.

Q. In Winitz' program do the subjects comprehend the material or do they in fact merely memorize it?
A. I don't think it's a rote memory task - there are too many items to memorize. The subjects have to grasp an unfamiliar word or concept and be able to relate it to the picture when it shows up again.

Q. I think that Harris (Winitz) has indeed shown to some degree that comprehension takes place. I remember one meeting at which he had us working in Russian. It was just a matter of a short time before we were comprehending quite well.

Q. In Winitz' program, is that a recorded voice? Could that be used to explain depressed scores for the aphasic subjects?
A. Yes, it is a recorded voice, and there has been some evidence that recorded presentation can affect the performance. We did not want to do it live voice, however, because we had no one who could speak German, as well as wanting to keep it the same across subjects.

Q. I'm bothered by the quantitative vs. qualitative distinction, especially since you're trying to show qualitative differences with numbers. You have a finding that was interesting to me about the aphasic subjects' difficulty with the associative aspects; I wonder if you might want to call that a qualitative difference.

A. We just didn't plan it that way. We looked at qualitative differences as other people have done before, by looking at different structures; but I like that, and we want to go back and look at that in a little more detail. It was something that came up in looking at the data, but we didn't carry it out.

Q. Were these subjects enrolled in therapy at the time of the testing?

A. Probably half were and half wern't.

Q. I'm curious as to what their reaction was when they came in for treatment and they were being taught German.

A. The aphasic patients didn't say much about it. Some of the normal subjects were a little confused by it and asked me what more we were going to do.

Q. This approach implies that you consider aphasia a loss and you might take a teaching approach in aphasia treatment.

A. Winitz in developing the program assumed that language is learned through auditory stimulation. We used this program to assess comprehension by trying to eliminate some of the cues we felt might make it easier for the normals to perform well on such tasks. I don't see that we would attempt to "teach" to aphasic adults as you suggest.