A Longitudinal Study of Word Fluency Response Strategies in Two Aphasic Subjects

Carl A. Coelho
Gaylord Hospital, Wallingford, Connecticut

Michael L. Kimbarow
SUNY College at New Paltz, New Paltz, New York

Mary Boyle
St. John's University, Jamaica, New York

Traditionally, word fluency measures have been scored by simply counting the number of words produced and comparing the total to some established value to determine if the number of words produced is comparable to similarly brain-damaged persons or normal controls. Recent investigations of fluency in controlled association have explored the diagnostic utility of qualitatively analyzing word fluency responses. For example, Collins, McNeil, Lentz, Shubitowski, and Rosenbek (1984) studied aphasic and nonaphasic subjects' use of semantic chaining and phonemic cueing while giving words beginning with the letter "s." Collins et al. reported that their subjects did not demonstrate any consistent response strategy on this task. Furthermore, the qualitative analyses failed to differentiate among aphasia subtypes.

Adamovich and Henderson (1984) studied semantic association strategies and various phonemic association strategies used by their subjects to generate words beginning with the letters "f," "a" and "s." Adamovich and Henderson observed that all groups, normal and brain-damaged, performed similarly with regard to the strategies used most often to generate lists of words. These strategies all fell within the phonemic association category. Semantic association strategies accounted for only 14% of the responses of normal subjects and 9% of all responses produced by brain-damaged subjects.

The similarity of findings reported by Collins et al. and Adamovich and Henderson regarding the minimal use of semantic association strategies during production of letter restricted word lists suggests the possibility that the tasks themselves did not permit use of such strategies. That is, there may be no inherent organization of semantic categories within a letter restricted word list. As an alternative to using letter restricted word fluency tasks Coelho, Kimbarow, and Maulucci (1984) reported a study in which they compared responses of aphasic patients on a letter restricted word fluency task, as well as a task in which subjects produced the names of animals and foods. They noted that production of lists of animals and foods lead to the use of semantic association strategies. This appeared to be related to structuring of superordinate and ordinate categories, such as farm animals or house pets, or generation of semantically associated chains, such as peas and carrots.

Finally, the results reported by Coelho et al., Collins et al. and Adamovich and Henderson were all based on data gathered after only one administration of the word fluency tasks. We felt that multiple presentations of such tasks might more clearly elucidate subjects' semantic association response strategies over time.

The purpose of this study was: (1) to longitudinally monitor the use of response strategies by two aphasic subjects on three word fluency tasks; and (2) to assess the impact of training on the development of such strategies.

METHOD

Subjects. RB was a 79-year-old right-handed female who was 26 months post onset of a single unilateral left hemisphere CVA. She was a retired teacher who demonstrated a mild aphasia (PICA Overall score of 13.85, 88th percentile). CP was a 76-year-old retired machinest, right-handed male who was 12 months post onset of a single unilateral left hemisphere CVA. He was moderately aphasic (PICA overall score of 12.18, 66th percentile).

Word Fluency Tasks. Three word fluency tasks were presented during this study. In two, subjects were asked to name all the animals and all the foods they could think of, and in the third, to list words beginning with the letter "s." All three tasks were timed for 60 seconds.

Design. Due to differences in the subjects' degree of aphasia two single-subject studies were conducted. Both employed multiple baseline acros behaviors designs (McReynolds and Kearns, 1983). Four baseline sessions were conducted prior to the initiation of treatment, each consisting of the presentation of the 3 word fluency tasks in randomized order. Treatment session were conducted 3 to 5 times per week. All sessions were thirty minutes long. Training began on the foods task for RB and on the animals task for CP while baselines were maintained on the untrained tasks (animals and "s" words for RB, foods and "s" words for CP). Treatment was modeled after Divergent Semantic Intervention (Chapey, 1981). Treatment sessions began with the subject asked to name all the foods (or animals) they could think of. As wor were given prompts were provided such as "Can you think of any other vegetables?" or "What other animals would you see on a farm?" At the completion each treatment session baseline probes were taken in a randomized order on ea fo the 3 word fluency tasks. Treatment continued on the first task for 15 sessions for RB. For subject CP treatment phases were extended to 25 session because of his more significant aphasia. Following completion of the first treatment phase, treatment began on the second word fluency task.

Social Validation. A social validity procedure (Thompson and Byrne, 1984) was used to estimate treatment effectiveness. Two normal adults, aged 77 and 83, matched closely with the aphasic subjects for level of education, were presented the three word fluency tasks in randomized order on 4 separate occasions. Responses were recorded, analyzed, and standard deviations c alculated.

Analysis. Subjects' responses were transcribed and the following analyses and tabulations performed. (1) Total number of words produced per word fluency task. (2) Total number of semantic associations (i.e. when two consecutively produced words of different subcategories were linked associatively, for example, "meat" and "potatoes"). (3) Total number of phonemic associations (i.e., when two consecutively produced words were linked by one of the following phonemic relations: identical initial blend -- "store, stop",consistent variation of one vowel -- "sale, seal", or a rhyme -- "potato, tomato").

Reliability. Intra- and interjudge reliability for all data analyzed resulted in 98% and 93% point-to-point agreement respectively.

RESULTS

The principal finding of this study was the marked degree of response variability both subjects demonstrated in total number of words produced on each of the 3 word fluency tasks (see Figures 1 and 2). This variability was consistent throughout the training programs for both RB and CP. Because

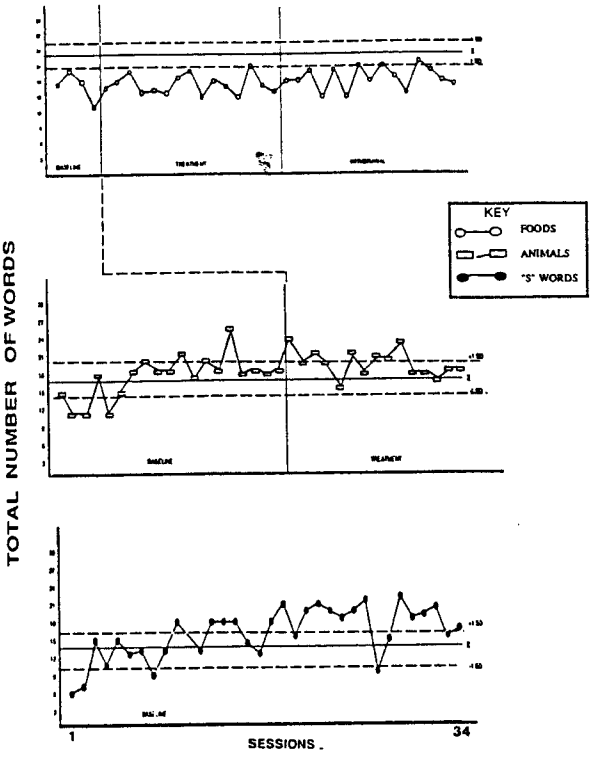

Figure 1. Total number of words produced by RB for foods, animals, and "S" words, and the mean and standard deviation for the social comparison data.

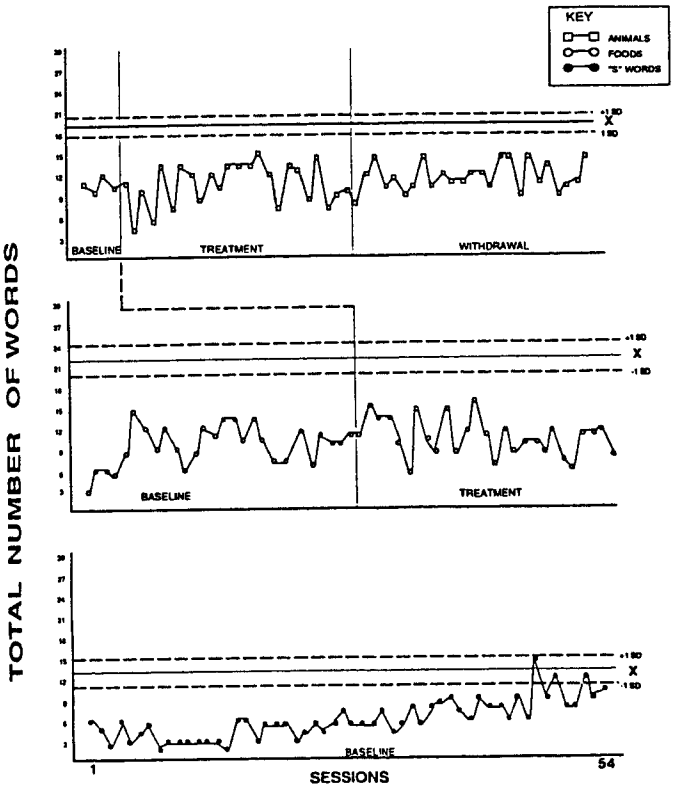

Figure 2. Total number of words produced by CP for animals, foods, and "S" words, and the mean and standard deviation for the social comparison data.

stable baselines were never established for either subject on any of the word fluency tasks, discussion of the effectiveness of training on the development of response strategies would not be meaningful. However, since both subjects were followed for extended periods of time (3 to 5 months) several observations can be made regarding both subjects' response patterns on these word fluency tasks.

Response Variability. Response variability was consistent for both subjects throughout the training program. The variability was at times quite dramatic. For example, RB produced 20 "s" words one day and only 9 the next. CP produced 11 animal names one day and 4 the following session.

This variability was noted not only in the responses of the aphasic subjects but also in those of the 2 normal adults used for the social validity procedure (see Figure 3). Stable baselines were obtained on only one of the word fluency tasks for one of the normal controls. On the other task the controls' performances appeared to improve with successive trials. This was most evident on the "s" words task in which 8 words were produced by one and 10 by the other during the initial session. By the fourth session they produced 19 and 21 words respectively.

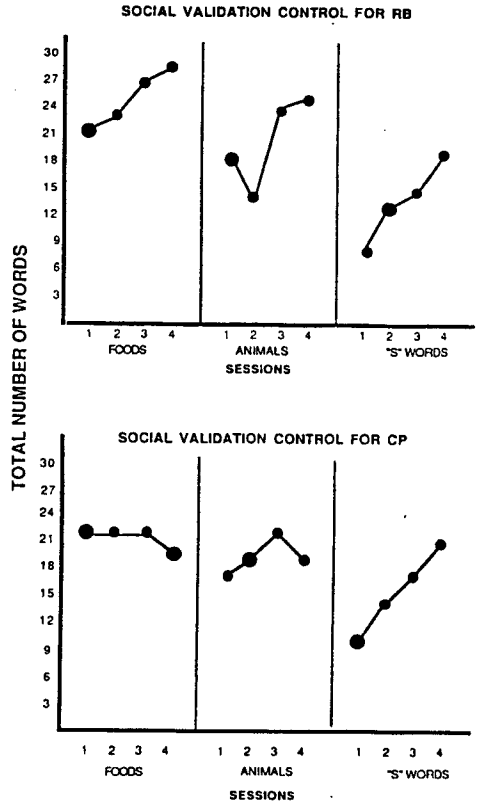

Figure 3. Total number of responses produced per session on three word fluency tasks by normal social validation subjects.

Differences Among Word Fluency Tasks. Both aphasic subjects were initially more fluent on the animals and foods tasks than on the "s" words task. Over the period of several sessions subject RB's (mildly aphasic)

range of fluency scores on the "s" words task gradually increased and intermittently approximated the ranges of scores for the animals and foods tasks. There were no consistent differences in the range of scores for the animals or foods tasks for either subject. These findings were inconsistent with those of Coelho et al. (1984) who noted that their aphasic subjects produced significantly more foods than animals and more animals than "s" words.

Response Strategies. Subject RB (mildly aphasic) demonstrated 236 instances of either semantic or phonemic strategy usage, while CP (moderately aphasic) made use of such strategies in only 57 instances. Semantic associations were the most frequently used response strategy by both subjects. Semantic strategies outnumbered phonemic strategies 96% to 4% for RB and 66% to 33% for CP. Both subjects demonstrated the most response strategy use with the "s" words task -- 80% of all instances for RB and 65% for CP occurred on this task. The most commonly used phonemic strategy by both subjects was the identical initial blend. The response strategies used by both RB and CP were predominantly self-generated and idiosyncratic. For example, RB, when confronted with the "s" words task consistently would visualize her bathroom and generate lists such as, "shower, sink, stopper, shampoo, soap, scrub, sponge, shoulder, scalp." RB's performance on the "s" words task improved over time and appeared to be related to her use of this semantic strategy. CP, on the other hand, had great difficulty with this task and would often take a dictionary and attempt to memorize several "s" words prior to each session. Figure 2 illustrates that this strategy was not at all productive for CP.

DISCUSSION

The findings of this study are based on just 2 subjects. However, data from these subjects were gathered longitudinally and therefore these findings have a variety of implications pertaining to the use of word fluency measures.

Response Variability. We were not surprised to see variability on our letter-restricted task. Wertz, Shubitowski, Dronkers, Lemme, and Deal (1985) have reported significant differences in the numbers of words produced in test-retest administrations within one week for letter-restricted tasks. Our results extend these findings to include semantically based tasks as well as a letter restricted task over several months time. The 2 aphasic subjects studied demonstrated marked session-to-session variability in the total number of words produced on each of the 3 word fluency tasks. Their performances varied by as many as 11 items on the same task over 2 consecutive sessions. The normal controls also demonstrated variability in their performances over 4 baseline sessions. Stable baselines were not established for the 2 normals on some of the word fluency tasks, and a practice/learning effect seemed to be present on the "s" words task, with the number of words produced steadily rising over time.

Word Fluency Tasks. Although the foods and animals tasks appeared to be easier for both aphasic subjects, in terms of their ability to produce greater numbers of words, the letter specific word fluency tasks are useful as well. The present findings support those of Coelho et al. (1984) that the foods and animals tasks lend themselves to superordinate and ordinate categorization by both aphasic and normal subjects. However, over time, semantic association strategies, which proved to be highly productive for subject RB and the normal controls, were noted to emerge on letter specific word fluency tasks. It may be that for the purposes of response strategy identification the foods and animals tasks are better suited for moderately aphasic subjects.

Response Strategies. Response strategies based on semantic association appeared to be productively used by both aphasic subjects. The consistency with which they were used, however, varied considerably between the two. This may have been a function of severity of aphasia. Both subjects typically initiated useful strategies to facilitate the listing of foods or animals. For example, on the foods task, RB regularly started with breakfast and progressed through the three meals, listing appropriate foods for each. CP would imagine he was in a grocery store pushing a shopping cart. However, it appeared as though the shelves were empty, as this strategy rarely yielded more than three items. CP would then switch strategies and attempt to list vegetables in his garden and eventually switch again to recalling what he had for lunch. CP's combination of strategies never yielded a greater number of words than RB produced using a single strategy.

Finally, in terms of treatment, semantic strategies were more easily facilitated than phonemic strategies with the aphasic subjects in the training program. The use of semantic strategies was easily incorporated into Chapey's (1981) Divergent Semantic Intervention technique.

Conclusions. The results of the present investigation permit the following statements regarding the use of word fluency measures.

1) The response variability for both aphasic and normal subjects on a variety of word fluency tasks suggests that serious consideration be given to presenting such tasks on more than one occasion and reporting performance on such tasks in terms of a range rather than a single score.

2) Tasks such as listing foods or animals appear to provide more information regarding response strategies in moderately aphasic subjects than letter specific tasks. However, letter specific tasks are equally useful with higher level aphasic individuals. It is probably best to use a "battery" of word fluency tasks to assess response strategies.

3) Qualitative analyses of word fluency responses can provide clinicians with far more information than simple tallies of words produced, particularly when these data are gathered longitudinally.

4) Using word fluency tasks as a measure of recovery from brain injury, as a basis for classifying subtypes of aphasia, or for the differential diagnosis of brain injury should be done cautiously.

REFERENCES

Adamovich, B. and Henderson, J. Can we learn more from word fluency measures with aphasic, right brain injured, and closed head trauma patients? In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1984. Minneapolis, MN: BRK Publishers, 124-131, 1984.

Chapey, R. Divergent semantic intervention. In R. Chapey (Ed.), Language Intervention Strategies in Adult Aphasia. Baltimore: Williams and Wilkins, 155-168, 1981.

Coelho, C., Kimbarow, M., and Maulucci, F. Word fluency measures in three groups of brain injured subjects. Paper presented at the annual convention of the American Speech-Language-Hearing Association, San Francisco, California, November 1984.

Collins, M., McNeil, M., Lentz, S., Shubitowski, Y., and Rosenbek, J. Word fluency and aphasia: Some linguistic and not-so-linguistic considerations. In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1984. Minneapolis, MN: BRK Publishers, Inc., 78-84, 1984.

McReynolds, L. and Kearns, K. Single Subject Experimental Designs in Communicative Disorders. Baltimore: University Park Press, 1983.

Thompson, C. and Byrne, M. Across setting generalization of social conventions in aphasia: An experimental analysis of "loose training." In R.H. Brookshire (Ed.), Clinical Aphasiology: Conference Proceedings, 1984. Minneapolis, MN: BRK Publishers, 132-144, 1984.

Wertz, R., Shubitowski, Y., Dronkers, N., Lemme, M., and Deal, J. Word fluency reliability in normal and brain damaged adults. Paper presented at the annual convention of the American Speech-Language-Hearing Association, Washington, D.C., November 1985.

DISCUSSION

Q: When you have a measure that is not reliable, you get significant differences from one administration to another without any intervention. Therefore, you probably have a measure that is not any good for measuring change as a result of intervention. I wonder if what we can say about the use of strategies isn't also affected by that. Can you really look at the development or use of strategies with an unreliable measure?

A: Probably not. I think we did, however, get some interesting findings in response patterns with the data being collected over several months.

Q: Were all of the subjects tested at the same time of the day across all the treatments?

A: Yes, typically late afternoons in their homes.

Q: Do you have any intuitions about why they are so variable on these kinds of tasks?

A: We really don't. We at one point thought it was the time of day and on a couple of occasions caught them early in the morning, but we saw no change in the variability.

Q: What kind of instructions did you give the subjects? Did you try to pump them up and get maximum performance?

A: After about five sessions the subjects knew what the intent of the task was. We always had some conversation to get them to relax but without coaching. I guess I did do some pumping up as well.

Q: Since normal people got so much better over time, and presumably normal people have stable language functions, it seems to me that this measure doesn't tell us anything about the status of a person's language function. Is there any reason clinicians continue to administer the measure?

A: I really don't have an answer for that. We are puzzled ourselves. One of the things I wish we had done was, using the WAB Classification System, plotted out the classification each subject would have fallen into based on his performance on any given day. They would have shifted back and forth between fluent and nonfluent.

C: It seems to me it doesn't tell us any more about their language skills than maybe having them counting the number of times they can tap their finger per minute.

A: I tend to use the measure more as a treatment observation to see what they can do with the task, particularly with head trauma patients.

C: I'd like to disagree with the previous comment. In fact, I think this sort of observation might be the fundamental that we're looking for to

try to find out why aphasics don't do things they can at other times do very well. So to say that it tells you nothing about language or nothing about how they use language is incorrect. The search for the variability might be the key that we are looking for.

C: It might be interesting and useful to combine this sort of study with priming studies to see whether the semantic fields of your subjects are changing or whether the problem is access. It seems clear that the semantic fields are largely intact, at least in most of these patients, but we don't know this on a longitudinal basis. If they are intact, it's not clear what the therapy really is addressing, why they should improve with semantically based therapy.

C: One of the points we wanted to make was that the measure was not appropriate for a diagnostic task. The earlier study that was alluded to was a paper we presented at ASHA a number of years ago. One of the striking differences that we found between our normal subjects and our brain damaged subjects, especially the aphasic subjects, was that the aphasic subjects typically would go down the supermarket shelves, pick out one or two items, list them, go up another aisle, take another two items, and back to the first aisle, take another two or three, then go back up the second aisle, maybe a third, and keep switching back and forth, whereas the normal subjects were listing all of the items within one aisle. They would exhaust the items, leave that aisle and go up the second one, exhaust all the items there, go down a third, and never go back. So one of the things we found strikingly different was that the aphasic subjects seemed to not have any opportunity to inhibit the exhaustion of a category, leaving it behind, and then going on to something else. So the semantic field boundaries seemed to break down somehow We were trying to figure out why that was going on, but we think that this is one of the most significant differences between normal function on this task and brain damaged function.

C: If you have a measure that is not reliable, you probably don't want to use it as a change measure. Fine, but I think this measure is telling us something. For example: if it is not reliable in normals and we find it is not reliable in brain damaged folks, then we learn something. If it were reliable in brain damaged folks and not in normals, we would have learned something else. There have been a few studies that start looking within the hemisphere, for intra-hemispheric differences on word fluency tasks. What Michael just pointed out is that perhaps we're looking at differences in problem solving on a task that simply uses verbal responses The difficulty may be in organizing or developing some kind of strategy that differs, depending on whether the lesion is anterior or prefrontal, compared to whether the lesion is posterior. For example, we know the prefrontal people just don't search very systematically on a trial and error nonverbal learning task, that once they find the solution they won't keep going back to it. They may get off and make the same kind of error they have made prior to finding the solution. Perhaps we are looking at something beyond just coming up with a certain number of words that begin with a letter or items in the supermarket. If they're running all over the supermarket, that's not a very efficient way to shop. Maybe we're not looking at language so much as the ability to formulate some kind of plan and reduce the active uncertainty in a problem.

A: I would agree completely. One of the differences we noted between the mild and moderate subject was that the mild subject seemed to make more consistent use of the strategy. They were both nonfluent types, but my intuition was that it was more a function of severity than type of aphasia, but perhaps that's another avenue to pursue.

C: I get bothered by saying we should throw out a task or we should not use a task because it is too variable when we don't really sort out where that variability is coming from. Specifically, I think if we are consistent in our administration of a task--whether it's word fluency or anything else--and we get a lot of variability in performance, then perhaps we shouldn't throw that task out. That may be a good task, and perhaps the variability is internal to the subjects. If you're consistent in your administration of that task and we get variability in performance, then maybe we can find out something about that subject's strategy or their variability.

Q: Could you tell me a little about your treatment? Specifically, did you train strategies, or did you give them tasks that were requiring them to be divergent and name things in these categories without a strategy?

A: Early in the project when we were planning the treatment phase, we had talked about trying to train strategies. If you will recall from the paper, I said many of the strategies they used were self-generated and idiosyncratic. We felt we might be stifling their use of a very productive self-generated strategy if we tried to impose our own strategy on them, so what we did was to identify the strategy they seemed to be using and then encourage its use by structuring the task. For example, the moderate aphasic seemed to do a lot of shifting; he would be going through the grocery store, and when that didn't work, he would go to what he had had for lunch, or what was in his garden. I would try to keep him from shifting out of a category until he had produced three or four items.

Q: Do you feel that if you had taken the "normal strategy," that is, marched them up the aisle and then down the second and the third that perhaps you may have gotten different results? It seems that if we take the idiosyncratic strategies and base our treatment on those as opposed to what we know about normal strategies (even though they are perhaps false and self-imposed), or what we see as a consistent strategy to access the information, those are two different approaches. I'm not sure one is better than the other, but my clinical hunch is that one leads in a more systematic direction. Rather than following the bouncing ball, so to speak, we might go the other way and follow the normals.

A: My feeling is that it is a question of severity. With the mildly aphasic subject I never would have attempted to train her to visualize her bathroom and then go through all the "s" words in her bathroom. I don't think I could do that myself, but it worked well for her.

Q: I wonder if you have thought about interhemispheric differences? The reason I said that is a couple of years ago Brenda Adamovich presented a paper looking at right hemisphere subjects on a number of tests, and one of them was word fluency. She observed, and we had observed also, that right hemisphere subjects don't seem to adopt any strategy and go from the grocery store to the garden to the fish market. If you looked at right hemisphere patients in the way you have looked at the lefts, you

might find something else out about what's going on and what the origins of their problems are. After all, these are people who have intact left hemispheres but are still not making normal associations.

A: In the earlier study we did look at right hemisphere subjects versus lefts and found that there were significantly more subcategory and within-category shifts with the rights than the lefts, so they were even more poorly organized in their response strategies than the lefts were.

Q: In the normal literature Battig and Montague and some other people have looked at a whole lot of these semantic categories and have found that for most of them there are certain words that most people tend to list right off. They are highly associated with that category and then they tend to move on to words not so highly associated with that category. When you move into this second echelon you start to have to use more deliberate strategies to help you remember. Do you have any idea as to whether your patients' response, if you look at them in those two ways, were more stable for the high association words than those which were harder to list? My idea is that maybe there would be more stability if you looked at just the high association lists, and then you get more variability only as you move into the others.

A: We haven't but that's a great idea, thank you.

C: We found, in our earlier study, that the right hemisphere group and the closed head injured group we evaluated started off their word lists with words that were probably fourth, fifth, or sixth level frequency-of-usage words. So, for example, on animals we would find the rights or head injured patients starting out, instead of with dog or cat, with aardvark or cobra, things we would probably get to after exhausting 15 or 20 items prior to that, so there seemed to be some breakdown in the way they were able to access lexicon to generate these word lists. We've never pursued that but it is something we have thought a lot about as to why that was going on, so we have some anecdotal data to suggest that is part of the difference between the interhemispheric groups.

C: It sounds like some people are getting some information from the word fluency tasks. But I'm still very concerned that clinicians use a change in score on word fluency to indicate a change in language use or comprehension or any language area. If normals are improving that much, and surely you don't think your normal subjects are improving in their ability to use language, then a change or improvement in a word fluency score doesn't necessarily tell you anything about the aphasic patient's change in language skills either.