

Connected Speech Characteristics of Right-hemisphere-damaged Adults: A Re-examination

Connie A. Tompkins, Richard Boada, Kathrine McGarry, Jennifer Jones, Angela E. Rahn, and Susan Ranier

Researchers have noted a number of deficiencies in speech samples elicited from adults with unilateral right hemisphere brain damage (RHD). These deficiencies include excessive literalness, verbosity, irrelevancies, overpersonalization, unnecessary or excessive detail and a tendency to itemize rather than interpret the events in an eliciting stimulus (Myers, 1986). However, the conclusions of studies undertaken to document and quantify these characteristics must often be interpreted with caution. Generally, the studies are limited by small sample sizes (Hobbs, Johnson-Emanuel, Molloy, & Tonkovich, 1989; Mackisack, Myers, & Duffy, 1987; Myers, 1979; Sherrad & Penn, 1990); selection biases such as all subjects receiving speech-language treatment, all exhibiting moderate-to-severe contralateral neglect, and/or all representing the acute post onset phase (Myers, 1979; Hillis Trupe & Hillis, 1985); and/or unreported reliability data (Hillis Trupe & Hillis, 1985; Hobbs et al., 1989; Myers, 1979).

The purpose of our study was to analyze connected speech samples elicited from a relatively large group of RHD patients who were not selected on the basis of severity or neurobehavioral status. Their samples were compared to those of two sociodemographically similar control groups: subjects with left hemisphere brain damage (LHD) and normally aging adults without known neurological impairment. Originally, the samples were collected as part of a larger study of nonliteral interpretation following brain damage (Tompkins, 1990). We intended to use the data to describe the speech characteristics of our three groups. As we analyzed the samples, we found that some of the features commonly thought to characterize RHD adults, such as low percentages of interpretive content units

(Myers, 1979), were not distinguishing our groups. Accordingly, we decided to pursue the analysis that is reported here. Following Myers' (1979) suggestion, we evaluated "unscorable content," which is not included in Yorkston and Beukelman's (1980) normative concepts list, by operationalizing a number of descriptors from the clinical literature.

METHOD

Subjects

This study tested 52 unilateral stroke patients (26 with RHD, 26 with LHD) and 26 controls without history of neurologic impairment. The patients had been consecutively admitted to two acute care settings and one rehabilitation center; they met study selection criteria and consented to participate. Patients were at least 4 months post onset of cerebrovascular accident (CVA), had no prior history of neurological disease, and had side lesions documented in CT scan reports. All subjects were right-handed, monolingual English speakers who passed a pure-tone hearing screening.

Table 1 summarizes subject characteristics for the three groups. The groups did not differ in age, education, estimated premorbid IQ, or gender distribution. Thirteen RHD subjects evidenced contralateral neglect in a complex figure-copying task, but in 12 cases it was mild (i.e., three or fewer elements of the design were omitted from the left side of the drawing). According to speech-language pathology records from their primary care facilities, fifteen of the LHD subjects were aphasic. More detail about screening criteria and clinical characteristics can be found in Tompkins (1990).

Stimulus and Task

As in many prior investigations of RHD adults' connected speech, the eliciting stimulus was the Cookie Theft picture from the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983). The picture was placed on a table in front of each subject, with these instructions: "I want you to look at this picture, and tell a story about everything you see happening."* The examiner gestured over the card to call attention to the

*These instructions vary slightly from the original instructions, which do not ask the speaker

TABLE 1. DEMOGRAPHIC AND CLINICAL CHARACTERISTICS OF THREE SUBJECT GROUPS

	<i>RHD</i> (<i>N</i> = 26)	<i>LHD</i> (<i>N</i> = 26)	<i>Control</i> (<i>N</i> = 26)
Age	62.9 (46-78)	64.8 (48-78)	62.7 (50-78)
Education (Years)	12.6 (8-20)	12.4 (8-20)	13.1 (8-20)
Gender	15 Male	16 Male	16 Male
Estimated (Premorbid) IQ	106 (84-124)	105 (89-128)	108 (94-127)
Months Post Onset of CVA	11.2 (4-20)	9.2 (4-16)	
Etiology of CVA	64% Thrombo- embolic	58% Thrombo- embolic	
• Auditory Comprehension (Overall BDAE %ile)	93 (83-98)	87 (28-98)	96 (91-98)
• Aphasia	100% No	34% No	100% No
• Judgment of Line Orientation (30 possible)	15.7 (0-29)	20.9 (8-29)	25.8 (20-30)
• Visual Neglect (Complex Figure Copy and Line Bisection Tasks)	64% No	96% No	100% No

Note: Data are means (with ranges) unless otherwise indicated.

Note: RHD denotes right-hemisphere-damaged; LHD denotes left-hemisphere-damaged

• indicates significant group differences ($p < .01$)

entire drawing. Speech samples were audiotape recorded and transcribed verbatim. An independent judge resolved any differences with the initial transcriber.

Scoring

Following procedures described in the primary references, the transcripts were scored for literal and interpretive content units (Myers, 1979; York-

to tell a story. This difference did not seem meaningful, as few subjects in any group attempted to produce a story. However, whether the same results would be obtained with slightly different instructions is an empirical question.

TABLE 2. "UNSCORABLE CONTENT" CATEGORIES FOR COOKIE THEFT SAMPLES

-
1. **Overpersonalization:** integration of self into story structure or description of events.
[Examples: "My son's home from school, making his snack raid on the cookie jar"; "if it was me, I'd open the window, throw the dishes out"; "this looks very much like my childhood, to tell you the truth"; "I could never imagine my mother being that young-looking"; "I've had days just like this"]
 2. **Excessive detail:** use of intensifiers; focus on tangential or minor elements of the picture.
[Examples: "The cookie jar has the words 'cookie jar' on it"; "she's industriously drying the dishes"; "it's a black-and-white line drawing"]
 3. **Value judgments:** judgments of a character's actions or imputed intentions.
[Examples: "That mother is irresponsible for not supervising her children more closely"; "he shouldn't be on that type of stepladder from the beginning, you know"]
-

ston & Beukelman, 1980), number of syllables (Yorkston & Beukelman, 1980), time in minutes, and phrase-length ratio (Goodglass, Quadfasel, & Timberlake, 1964). Phrase-length ratio (PLR) is a measure of speech fluency that weighs the number of uninterrupted speech runs that are five words or more in length against the number of one- and two-word utterances produced. Samples were independently scored by two raters; differences were resolved with a third judge when necessary. Four samples selected at random from each group were re-scored after a period of at least two weeks with at least 90% agreement. More detail about reliability standards and procedures is provided in the Appendix.

Following the suggestion of Myers (1979), we also evaluated "unscorable content"—information not captured in Yorkston and Beukelman's (1980) normative list. Two scorers independently coded each transcript for unscorable content, using categories such as overpersonalization, excessive detail, value judgments, related errors, asides, and picture-related literal and interpretive concepts. A third rater, blind to subject group, also classified the samples. We achieved at least 88% scoring agreement for three categories: overpersonalizations, excessive detail, and value judgments. Table 2 provides the definitions and examples for each of these categories.

When one of these categories was coded, raters also noted the extent to which that category was represented throughout each transcript. A "Low Occurrence" was indicated for a particular category when one or two examples from that category were present in the sample, and a "High Occurrence" rating was given when three or more examples from that category were identified in the transcript. This High versus Low rating for each category was made with 100% agreement. The other distinctions that

we wished to code were redefined and re-analyzed several times but ultimately were dropped due to insufficient reliability.

RESULTS

Group mean data were calculated first for numbers of literal and interpretive concepts, percent literal and interpretive concepts, time in minutes, number of syllables, efficiency measures (number of literal and interpretive concepts produced per minute, and rate), and PLR (see Table 3). Only two of these measures distinguished the groups: rate, in syllables per minute [$F(2, 75) = 9.8; p < .01$] and PLR [$F(2, 75) = 5.4; p < .01$]. Post-hoc tests of means using the Tukey A procedure indicated that RHD and control groups did not differ, but that LHD subjects had slower speaking rates and smaller PLRs (indicating more hesitant speech) than the other two groups.

The number of subjects in each group who produced High or Low levels of each of the unscorable content categories was determined (see Table 4). Unscorable content occurred infrequently, regardless of group. Because of the small expected frequencies in each cell, High versus Low occurrence data were combined within each group prior to data analysis (Siegel, 1956). The frequency of unscorable content did not distinguish the three groups' samples [$\chi^2(4) = 1.81; p > .05$].

A variety of factors were inspected to ascertain their association with the production of unscorable content, including age, education, estimated (premorbid) IQ, gender, PLR, percent literal concepts, number of syllables produced, visual neglect for RHD subjects, and aphasia and auditory comprehension level for LHD subjects. Within each group, subjects who did or did not generate unscorable content were differentiated by the number of syllables in their transcripts; when any of the three types of unscorable content occurred, more syllables were uttered (all $t > -3.2; p < .005$). There were no other distinguishing measures within normal control and RHD groups, but the results for neglect deserve comment. The presence of neglect was not associated with the frequency or occurrence of unscorable content in our RHD group [$\chi^2(1) = 0.01, p > .05$], but this finding must be interpreted with caution as the neglect was mild in all but one case. Within the LHD group, proportionately fewer subjects who produced unscorable content were aphasic (33% of the subgroup) than subjects who did not produce unscorable content (66% of the subgroup), but this difference was not significant ($\chi^2(1) = 1.5; p > .05$). Auditory comprehension level and PLR (speech fluency) also did not distinguish subjects within the LHD group.

TABLE 3. GROUP MEAN DATA FOR QUANTIFYING PICTURE DESCRIPTIONS

	<i>RHD</i>	<i>LHD</i>	<i>Control</i>
Number of Literal Concepts			
Mean	9.7	8.0	9.6
(S.D.)	(3.8)	(3.7)	(3.4)
Range	4-22	0-18	4-18
Number of Interpretive Concepts			
Mean	5.6	5.6	6.3
(S.D.)	(2.1)	(2.8)	(2.3)
Range	2-8	0-11	2-10
Percent Literal Concepts			
Mean	64	60	60
(S.D.)	(12)	(14)	(8)
Range	44-87	31-100	42-75
Time in Minutes			
Mean	0.8	1.1	0.8
(S.D.)	(.6)	(.7)	(.3)
Range	0.3-2.9	0.3-3.3	0.3-1.4
Literal Concepts Per Minute			
Mean	13.1	10.7	14.0
(S.D.)	(6.6)	(8.0)	(5.0)
Range	4-27	1-40	4-27
Interpretive Concepts Per Minute			
Mean	7.0	7.8	8.2
(S.D.)	(5.9)	(4.4)	(4.2)
Range	1-27	2-18	3-18
Number of Syllables			
Mean	133	111	135
(S.D.)	(82)	(70)	(56)
Range	56-470	23-319	44-253
Rate (Syllables/Minute)**			
Mean	166	121	177
(S.D.)	(53)	(55)	(44)
Range	64-247	7-228	89-242
Phrase-length Ratio (PLR)**			
Mean	4.2	2.0	4.0
(S.D.)	(2.9)	(2.1)	(2.8)
Range	0.8-12	0.0-9.0	0.4-10.0

**significant, $p < .01$; LHD different from other two groups.

TABLE 4. NUMBERS OF SUBJECTS PRODUCING RESPONSES IN THREE "UNSCORABLE CONTENT" CATEGORIES

	<i>RHD</i>	<i>LHD</i>	<i>Control</i>
Overpersonalization			
High Extent	2	0	1
Low Extent	1	1	3
Value Judgments			
High Extent	0	1	1
Low Extent	2	2	1
Excessive Detail			
High Extent	1	1	3
Low Extent	4	4	4

DISCUSSION AND IMPLICATIONS

The principal result of these analyses is that some of the disordered behavior ascribed to RHD subjects, such as high proportions of literal concepts, overpersonalization, and excessive detail, did not distinguish RHD subjects from LHD or normally aging control subjects. Generalizing findings from the most severely impaired RHD patients (e.g., those seen in treatment, those in an acute post-CVA stage, or those with marked contralateral neglect) to the entire population is not warranted. For purposes of replication and generalization, it is critical that researchers provide explicit information about the way in which their subjects were sampled. In addition, clinical and/or behavioral indicators of severity that may be important for assessing individual variability need to be assessed routinely (see Bloise & Tompkins, 1993).

The fact that the analysis of unscorable content did not distinguish between groups was not entirely surprising. As indicated above, the generality of the findings from several of the earlier studies was questionable, due to various selection biases. In addition, Mackisack et al. (1987) also reported some similarities between RHD and control samples in a related analysis. But because explicit statements limiting generality are infrequent, we suspected that many readers would tend to overgeneralize from the earlier literature, without recognizing the heterogeneity of RHD subjects (see also Joannette, Goulet, & Hannequin, 1990).

To follow up on our impression about overgeneralization, we asked 10 speech-language pathologists with a range of 1 to 20 years' experience in rehabilitation settings to classify speech samples from the RHD and control groups. They sorted the transcripts into two groups and provided

reasons for their decisions. All clinicians were at chance levels in categorizing the normal samples. They judged normally aging speakers to be RHD principally on the basis of two characteristics that we studied and that have been emphasized in prior literature: excessive detail and over-personalization. These results dovetail with the present analyses and indicate the need for a more careful approach to defining and attributing "disordered" communication status (Ranier, Tompkins, Boada, & Spencer, 1991).

There are several limitations of our speech sample analysis that should be emphasized here. First, our coding system may not have included some important characteristics. In certain cases, we had the impression that there was something qualitatively unusual about the RHD samples that our analyses did not capture. For instance, we tried to code some of the unusual features as visual errors, but coders blind to subject group could not distinguish visual errors from word-retrieval errors with sufficient reliability. Related to the issue of coding schemes, Mackisack, Myers, & Duffy (1987) reported that verbosity and labelling behavior were more evident in an RHD group when tangential information, irrelevancies, and excessive detail were removed from their speech samples. In any case, different behavior-coding categories should be examined in further investigations of the connected speech of RHD adults.

The restrictive nature of our single discourse task also limits our results. Some suggest that procedural discourse tasks may be more sensitive for detecting the disorganization and problems of coherence that are frequently observed in the discourse of RHD adults (Sherrad & Penn, 1990; Ulatowska, H. K., personal communication, June 1991). Shadden, Burnette, Eikenberry, and DiBrezzo (1991) reminded us of an assortment of attributes and processing demands inherent in eliciting procedures that contribute to patterns of discourse production in normally aging adults. Information from a variety of language sampling contexts will be required to obtain a fuller picture of the discourse production abilities of RHD patients.

As a concluding point, our experience with these samples suggests that the content unit analysis of Cookie Theft descriptions may also require some revision. Our control subjects produced a number of new concepts that could not be scored using Yorkston & Beukelman's (1980) normative analysis and that did not fit into our coding categories, such as comments about the weather, the apron that the mother is wearing, and the yard and shrubs. For clinical purposes, we are expanding the content unit list to include these new concepts, as well as regionalisms and synonyms. If other clinicians developed regional lists, it might be possible to integrate them to achieve an expanded national normative list. In the meantime, our regional list will be available on request to anyone who is interested.

ACKNOWLEDGMENTS

This manuscript was supported by grant #DC00453 from the National Institute on Deafness and Other Communication Disorders.

REFERENCES

- Bloise, C. G. R., & Tompkins, C. A. (1993). Right brain damage and inference revision, revisited. In M. L. Lemme (Ed.), *Clinical Aphasiology* (Vol. 21, pp. 145-155). Austin, TX: PRO-ED.
- Goodglass, H., & Kaplan, E. (1983). *The Boston Diagnostic Aphasia Examination* (2nd ed.). Philadelphia, PA: Lea & Febiger.
- Goodglass, H., Quadfasel, F. A., & Timberlake, W. H. (1964). Phrase length and the type and severity of aphasia. *Cortex*, 1, 133-153.
- Hobbs, P. A., Johnson-Emanuel, C., Molloy, J. L., & Tonkovich, J. D. (1989, November). The diagnostic utility of the Cookie Theft picture relative to right hemisphere damage. Poster session presented at the American Speech-Language-Hearing Association annual convention, St. Louis, MO.
- Hillis Trupe, E. & Hillis, A. (1985). Paucity vs. verbosity: Another analysis of right hemisphere communication deficits. In R. H. Brookshire (Ed.), *Clinical Aphasiology* (Vol. 15, pp. 83-96). Minneapolis, MN: BRK Publishers.
- Joanette, Y., Goulet, P., & Hannequin, D. (1990). *Right hemisphere and verbal communication*. New York: Springer-Verlag.
- Mackisack, E. L., Myers, P. S., & Duffy, J. R. (1987). Verbosity and labeling behavior: The performance of right hemisphere and non-brain-damaged adults on an inferential picture description task. In R. H. Brookshire (Ed.), *Clinical Aphasiology* (Vol. 17, pp. 143-151). Minneapolis, MN: BRK Publishers.
- Myers, P. S. (1979). Profiles of communication deficits in patients with right cerebral hemisphere damage. In R. H. Brookshire (Ed.), *Clinical aphasiology conference proceedings* (pp. 38-46). Minneapolis, MN: BRK Publishers.
- Myers, P. S. (1986). Right hemisphere communication impairment. In R. Chapey (Ed.), *Language intervention strategies in adult aphasia* (2nd ed.) (pp. 444-461). Baltimore, MD: Williams and Wilkins.
- Kanier, S., Tompkins, C. A., Boada, R., & Spencer, K. A. (1991 November). *Connected speech in aging: Some cautions when attributing abnormality*. Paper presented at the American Speech-Language-Hearing Association annual convention.
- Shadden, B., Burnette, R., Eikenberry, B., & DiBrezzo, R. (1991). All discourse tasks are not created equal. In T. E. Prescott (Ed.), *Clinical Aphasiology* (Vol. 20, pp. 327-342). Austin, TX: PRO-ED.
- Sherrad, S. M. & Penn, C. (1990). Discourse in a right-hemisphere brain-damaged subject. *Aphasiology*, 4, 539-560.
- Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw-Hill.
- Tompkins, C. A. (1990). Knowledge and strategies for processing lexical metaphor after right or left hemisphere brain damage. *Journal of Speech and Hearing Research*, 33, 307-316.
- Yorkston, K. M. & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of Speech and Hearing Disorders*, 45, 27-36.

APPENDIX PROCEDURES AND STANDARDS FOR ASSESSING INTER- AND INTRA-SCORER RELIABILITY

Number of literal concepts: Agreement within plus or minus one concept was achieved for all samples evaluated.

Number of interpretive concepts: Agreement within plus or minus one concept was achieved for all samples evaluated.

Time in minutes: Time was measured with a stopwatch directly from the audiotape; agreement within plus or minus three seconds was achieved for all samples evaluated.

Number of syllables: All syllables (including interjections, unintelligible syllables, for example) were counted from transcriptions; agreement within plus or minus 5% was achieved for all samples evaluated.

Phrase-length ratio: This measure involves tallying the number of uninterrupted units of 5 words or more (numerator of the ratio), and the number of 1- or 2-word uninterrupted units (totalled in the denominator of the ratio). These tallies were taken from transcripts that included notations for pauses and other interruptions. Agreement for each component of the ratio was within plus or minus one unit for all samples evaluated.