Journal of Speech and Hearing Research, Volume 38, 145–156, February 1995

Presence, Completeness, and Accuracy of Main Concepts in the Connected Speech of Non-Brain-Damaged Adults and Adults With Aphasia

Linda E. Nicholas
Robert H. Brookshire
Minneapolis Veterans Affairs Medical Center
and
University of Minnesota, Minneapolis

A standard rule-based system was used to evaluate the presence, accuracy, and completeness of main concepts in the connected speech of 20 non-brain-damaged adults and 20 adults with aphasia. Main concepts form a skeletal outline of the most important information (or "gist") in a message. The interjudge and intrajudge reliability of the main concept scoring system and the test-retest stability of scores were acceptable. The non-brain-damaged group produced significantly more *Accurate/complete* main concepts, and significantly fewer *Accurate/incomplete*, *Inaccurate*, and *Absent* main concepts than the group with aphasia. However, when the performance of individual subjects was evaluated, what best discriminated the performance of subjects with aphasia from that of non-brain-damaged subjects was not the number of main concepts they failed to mention but the accuracy and completeness of the main concepts they did produce. Measures of main concept production may be a clinically useful complement to other measures of communicative informativeness and efficiency.

KEY WORDS: aphasia, discourse, concepts, adults, scoring

Various measures have been used to compare the connected speech of adults with aphasia to that of non-brain-damaged adults and to evaluate changes in connected speech over time. These measures span a continuum from those used to assess adherence to standard language rules and patterns of use to those used to evaluate the informativeness and efficiency of connected speech. Measures of adherence to standard language rules and patterns of use include *counts of syntactic errors* (Shewan, 1988; Wagenaar, Snow, & Prins, 1975), *ratio of clauses to terminal units* (T-units) (Hunt, 1965), *type-token ratio*, *mean length of utterance*, and *number and types of cohesive ties* (Halliday & Hasan, 1976). Measures of communicative informativeness and efficiency include *content units per minute* (Yorkston & Beukelman, 1980), *percent of words that are correct information units* (Nicholas & Brookshire, 1993), and *subjective ratings of coherence* (Ulatowska, Freedman-Stern, Doyel, & Macaluso-Haynes, 1983).

Which measures will be most useful for clinically evaluating the connected speech of adults with aphasia may depend on the severity of an individual's aphasia and his or her communication needs. Measures of grammatical complexity and accuracy, utterance length, lexical diversity, and cohesion may provide important information about aphasic adults with mild language impairments who may be returning to work or participating in other situations requiring connected speech that adheres to standard language rules and patterns of use. For aphasic adults with greater

0022-4685/95/3801-0145

language impairments and whose communicative interactions are less demanding, the emphasis in assessment and treatment may be on how effectively and efficiently they communicate the content of messages to listeners, rather than on the form of the messages. Although their connected speech may not adhere to standard language rules or patterns and may require greater than normal comprehension effort on a listener's part, it still may effectively communicate their message.

Several standard measures for evaluating the communicative informativeness and efficiency of aphasic adults' connected speech have been published. Two such measures—*content units* (Yorkston & Beukelman, 1980), and *correct information units* (CIUs) (Nicholas & Brookshire, 1993)—quantify the number of informative words spoken.[1] The number of informative words may be evaluated in relation to the total words spoken (percent CIUs) or to the rate at which the informative words are spoken (content units per minute and CIUs per minute). Although these measures quantify several aspects of communicative informativeness and efficiency, they fail to measure other aspects, such as how much of the main information or gist about a topic a speaker conveys.

The literature suggests that speakers establish the main points in discourse by selectively highlighting some information by means of repetition and elaboration, and by establishing referential and causal connections between units of information (Kintsch & van Dijk, 1978; Schank & Abelson, 1977; van Dijk & Kintsch, 1983). This highlighted information forms a skeletal outline of the discourse, which van Dijk (1980) has labeled the *macrostructure*. According to van Dijk (1985), the macrostructure of a discourse "explains what is most relevant, important, or prominent in the semantic information of the discourse as a whole" (p. 115) and "is the semantic information that provides overall unity to the discourse" (p. 116). Kintsch and van Dijk (1978) developed the concept of macrostructure during formulation and evaluation of models of discourse comprehension. They reported that information from discourse that is closely associated with the macrostructure (the main ideas) is more likely to be comprehended and retained than peripheral information (the details). Many other studies also have found that the main information in discourse is comprehended and retained better than the details (Johnson, 1970; Meyer, 1975; Meyer & McConkie, 1973; and others).

Because the salience of information in the overall structure of discourse has strong effects on listeners' comprehension, it seems likely that aphasic adults' success in conveying the main information about a topic will be positively related to their overall communicative success. How well brain-damaged speakers convey the main information about a topic has been evaluated in the picture descriptions and story retellings of adults with aphasia (Gleason et al., 1980; Ulatowska, et al., 1983), aphasia or dementia (Heir, Hagenlocker, & Shindler, 1985; Nicholas, Obler, Albert, & Helm-Estabrooks, 1985), and right brain damage (Joanette, Goulet, Ska, & Nespoulous, 1986). The results of these studies suggest that, when adults with brain damage are considered as a group, they are likely to produce fewer main units of information than their non-brain-damaged counterparts. However, adults with aphasia may give special attention to main information when they generate discourse. Ulatowska and colleagues (1983) noted that the narratives of the adults with aphasia in their study included essential propositions (main units of information) more frequently than other propositions. In these five studies, main units of information were scored simply as *present* or *absent*—their accuracy and completeness were not evaluated. Additionally, the procedures used for determining the main units of information and for scoring their absence or presence are not described in sufficient detail to allow use of these procedures by others. Finally, the authors of these studies do not report the intrajudge or interjudge reliability or test-retest stability of their scoring procedures.

In this report, we will:

1. Describe procedures for determining main concepts for speech elicitation stimuli and for scoring their presence, completeness, and accuracy in connected speech.[2]
2. Provide intrajudge and interjudge reliability data for the main concepts scoring procedure.
3. Provide test-retest stability data for main concepts scores.
4. Compare the presence, completeness, and accuracy of main concepts in the connected speech of adults with aphasia to that of non-brain-damaged adults.

Method

Subjects

Subjects were 20 non-brain-damaged adults (10 male, 10 female) and 20 adults with aphasia (18 male, 2 female). All were right-handed native speakers of English. Aphasic subjects were recruited from past and current caseloads of speech-language pathology clinics in the Minneapolis-St. Paul metropolitan area. Each was at least 3 months post onset of a single left-hemisphere thromboembolic brain injury. They were diagnosed as aphasic by speech-language pathologists based on results of standard tests. The diagnosis of aphasia was subsequently confirmed by the first author, using standard test results and observation of the subjects in testing and in communicative interactions. To be diagnosed as aphasic, a subject had to exhibit a language-specific multimodality deficit. The aphasic subjects exhibited a range of type and severity of aphasia. The severity of their aphasia was estimated by their overall percentile on a four-subtest shortened version (SPICA) (Disimoni, Keith, & Darley, 1980) of the Porch Index of Communicative Ability

[1]Yorkston and Beukelman measured information in terms of "content units," which they defined as "a grouping of information" that was always expressed as a unit by at least 1 of their 78 normal speakers (p. 30). Their content units ranged from one word to several words in length.

[2]We have chosen the term "main concepts" to label this aspect of connected speech. *The Random House Dictionary of the English Language* (Stein & Urdang, 1973) defines "concept" as "a general notion or idea; an idea of something formed by mentally combining all its characteristics or particulars; a construct."

(Porch, 1971). All aphasic subjects' overall SPICA scores were at least two standard deviations below the overall PICA mean for non-brain-damaged adults (Duffy, Keith, Shane, & Podraza, 1976). The aphasic subjects also were tested with the Boston Naming Test (BNT) (Kaplan, Goodglass, & Weintraub, 1983), using revised administration and scoring procedures described by Nicholas, Brookshire, MacLennan, Schumacher, and Porrazzo (1989). They ranged in age from 51 to 77 years, and in education from 10 to 16 years. Descriptive information and test scores for the aphasic subjects are given in Table 1.

Non-brain-damaged subjects were nonhospitalized adults who lived independently and reported no history of speech, language, neurologic, or psychiatric problems. None showed evidence of cognitive, speech, or language impairments during the pretests and practice described subsequently. They ranged in age from 50 to 73 years ($M = 64.2$; $SD = 7.0$) and in years of education from 8 to 16 ($M = 12.8$; $SD = 2.2$).

Stimulus Materials

Connected speech was elicited with 10 stimuli, all of which were likely to be familiar to most North American adults. Four were single pictures, two were picture sequences, two were requests for personal information, and two were requests for procedural information. Two of the single pictures were from standard aphasia tests (the "cookie theft" picture from the Boston Diagnostic Aphasia Examination [BDAE; Goodglass & Kaplan, 1983] and the "picnic" picture from the Western Aphasia Battery [WAB; Kertesz, 1982]). The other two single pictures and the two picture sequences were drawn to the authors' specifications by a professional artist. (These pictures are shown in Nicholas & Brookshire, 1993.) The two single pictures each depicted a story-like situation with a central focus and interactions among pictured elements. Each implied a series of events leading up to the pictured scene, and each suggested events that were likely to follow the pictured situation. The two picture sequences each contained six pictures that related a story. The requests for personal information (Tell me what you usually do on Sundays. Tell me where you live and describe it to me.) and the requests for procedural information (Tell me how you would go about doing dishes by hand. Tell me how you would go about writing and sending a letter.) were selected from larger sets of requests based on the results of a pilot study, which suggested that they would elicit speech samples with a reasonable amount and consistency of content across speakers.

TABLE 1. Descriptive information for 20 aphasic subjects.

	Age	Education	Months PO	SPICA %ile	BNT
Mean	64.9	13.1	56.2	63.7	31.0
SD	6.8	1.7	62.7	14.5	12.1
Range	51–77	10–16	3–192	40–85	0–48

Note. Months PO = months post-onset of aphasia. SPICA %ile = overall percentile on a four-subtest version of the Porch Index of Communicative Ability. BNT = number correct of 60 possible on the Boston Naming Test.

Procedures

All testing sessions were conducted in a quiet room, free from distractions. The subject and the examiner sat side-by-side at a table on which was an audiocassette recorder and a microphone. Each subject's spoken responses to the elicitation stimuli, together with any prompts delivered by the examiner, were recorded.

Pretest and practice. To ensure that subjects had adequate vision for the task, each subject was asked to match line drawings of single objects to the same objects in a composite line drawing (the speech elicitation picture from the Minnesota Test for Differential Diagnosis of Aphasia [Schuell, 1972]). Subjects were shown, one at a time, six single objects from the composite picture and asked to point to the same object in the composite picture. Subjects had to match correctly all six to participate in the study. No subject failed to pass this screening test. Subjects' hearing and comprehension of spoken instructions was assessed subjectively during the pretests and practice. All subjects were judged to have adequate hearing and comprehension for the study tasks.

To ensure that subjects' speech was sufficiently intelligible to permit accurate transcription and that subjects produced enough speech to permit meaningful scoring of content, each subject was asked to describe a three-picture sequence that related a short story. In order to participate in the study, a subject had to produce at least 10 intelligible, relevant, nonrepeated words in response to the picture sequence. If a subject failed to meet the criterion for inclusion on a first try, the subject was given a second (and final) chance to describe the same picture sequence. Two potential subjects who exhibited severe aphasia were unable to meet this criterion and were excluded from the study.

The 40 subjects who qualified for the study were given a short interval of practice and training with two stimuli that were not used in the study (Tell me what you like to spend your time doing. Tell me how you would go about making a sandwich.). Instruction and feedback were provided, as needed, until the examiner felt that the subject understood the task. Feedback and instruction generally took the form of letting the subject know if his or her response was considered satisfactory in length and content. A response was considered satisfactory if it lasted for at least 15 sec and contained some information that was relevant to the stimulus. If it was not considered satisfactory, the subject was given suggestions about other information that could have been provided. No subject was given more than two trials and associated training with the practice stimuli.

Elicitation of speech samples. The stimulus pictures and spoken requests were presented individually to subjects in random order. When the eliciting stimulus was a picture or a picture sequence, it was placed on the table in front of the subject when the examiner asked the subject to talk about it, and was left there until the subject finished responding. Subjects were instructed to tell what they saw happening in

the picture(s). When the eliciting stimulus was a spoken request for personal or procedural information, the examiner simultaneously spoke the request and placed a printed request on the table in front of the subject. The printed request, which was typed on a card in bold 1/4″ high letters, was left on the table until the subject finished responding. Subjects were asked to try to talk about each stimulus for approximately 1 minute. If a subject stopped talking before producing at least 15 seconds of speech, he or she was prompted once with "Can you tell me more?" No further prompts were given. The examiner provided no feedback regarding the accuracy or appropriateness of subjects' responses, but did provide occasional social continuants such as "uh-huh" and head nods.

A different random order of stimuli was established for each subject, and each subject responded to the 10 stimuli in the same order in three sessions. The subjects were told at the beginning of the first session that they would be asked to talk about the stimuli three times. The first two sessions took place on the same day and were separated by a 10-minute break. The third session took place 7 to 10 days following the first two sessions. Practice was provided only at the beginning of the first session.

Transcribing speech samples. Subjects' recorded speech samples were orthographically transcribed by a speech-language pathologist familiar with the speech of adults with aphasia. A second speech-language pathologist independently checked the transcriptions against the tapes. Transcription disagreements between the two were resolved by the first author.

Validation of main concepts. We were unable to determine main concepts for the two personal information stimuli (Tell me what you usually do on Sundays. Tell me where you live and describe it to me.) because Sunday activities and residence descriptions were in large part idiosyncratic to individual subjects. Consequently no consistent set of concepts was given by validation subjects for either of these two stimuli. Main concepts were determined for the other 8 elicitation stimuli through a two-stage process. In the first stage, 10 speech-language pathologists read rules for writing main concept statements and participated in several training tasks. They were told that their main concept statements should provide an outline of the gist or essential information portrayed in the stimulus pictures, or an outline of the essential steps in the procedures. They were also told that their main concept statements should contain one and only one main verb. Following feedback on their performance in the training tasks, they were asked to write a list of main concepts for each of the eight elicitation stimuli. Main concepts that were listed by at least 7 of the 10 judges were put on a preliminary list. This list contained 64 main concepts (Table 2).

In the second stage, the first author and another speech-language pathologist used this preliminary list of main concepts along with written scoring rules to identify and score main concepts in the transcripts of the 20 non-brain-damaged subjects (one session per subject). After main concepts were identified, the two scorers assigned one of the following four scores to each: Accurate/complete (AC); Accurate/incomplete (AI); Inaccurate (IN)[3]; or Absent (AB).

For validation purposes, main concepts that received AC, AI, or IN scores were considered to be present in the transcript. Point-to-point interjudge reliability for determining the presence or absence of main concept statements was calculated on the transcripts of 12 randomly selected subjects. Reliability exceeded 90% for all 12 subjects (M = 97%). Main concepts that were present in the transcripts of at least 14 of the 20 non-brain-damaged subjects were put on the final main concept lists. The final lists for the 8 stimuli contained 53 of the original 64 main concepts (Table 2). No additional concepts that were not on the preliminary main concept lists were mentioned by 14 or more of the 20 non-brain-damaged subjects.

Scoring transcripts. First, the two individuals who participated in the validation study practiced scoring main concepts using the transcripts of 6 subjects (one session each). Then they compared their scoring decisions, discussed disagreements, and clarified misunderstandings. Following this practice scoring, they each independently scored main concepts in half of the remaining transcripts using the four scoring categories listed above. Written scoring rules (Appendix A) and the main concept lists and scoring examples were used to guide their scoring decisions. (A sample list of main concepts and scoring examples for these main concepts are given in Appendix B.)

Results

Reliability of Scoring

To assess interjudge reliability, 8 speech samples from one randomly selected session for each of 12 non-brain-damaged and 12 aphasic subjects were selected. The aphasic subjects were systematically selected to represent the group in severity of aphasia. Their SPICA overall percentiles ranged from 40 to 85 (M = 67.9). The two original scorers then each independently scored those transcripts previously scored by the other. Point-to-point interjudge reliability was calculated with the following formula:

$$[\text{Total agreements}/(\text{Total agreements} + \text{Total disagreements})] \times 100.$$

Interjudge reliability for each of the four scoring categories exceeded 80 percent (M = 86%). Interjudge reliability was 80% or higher for all 12 non-brain-damaged subjects (M = 94%) and for 11 of the 12 aphasic subjects (M = 86%). Reliability was 74% for the most impaired aphasic subject in the study (SPICA OA percentile = 40).

Interjudge reliability also was calculated between the first author's scores and those of another speech-language pa-

[3]In the first version of the scoring categories, the Inaccurate category was divided into two subcategories—Inaccurate/complete and Inaccurate/incomplete. However, we were unable to establish adequate interjudge reliability for the Inaccurate/incomplete category, because it was difficult to judge the completeness of main concepts that contained inaccuracies.

TABLE 2. The number (and percent) of 10 speech-language pathologists (SLPs) who mentioned main concepts for each of eight elicitation stimuli, and the number (and percent) of 20 non-brain-damaged (NBD) subjects who also mentioned those main concepts.

Stimulus	Group	1	2	3	4	5	6	7	8	9	10
BDAE	SLP	10 (100)	10 (100)	9 (90)	10 (100)	10 (100)	10 (100)	7 (70)			
	NBD	18 (90)	20 (100)	15 (75)	20 (100)	17 (85)	15 (75)	15 (75)			
WAB	SLP	8 (80)	7 (70)	7 (70)*	10 (100)	7 (70)	10 (100)	10 (100)	10 (100)		
	NBD	19 (95)	16 (80)	12 (60)*	19 (95)	15 (75)	17 (85)	18 (90)	16 (80)		
Birthday	SLP	9 (90)	10 (100)	9 (90)	7 (70)	10 (100)*	10 (100)	9 (90)			
	NBD	19 (95)	20 (100)	20 (100)	17 (85)	12 (60)*	14 (70)	20 (100)			
Cat	SLP	10 (100)	10 (100)	7 (70)*	9 (90)	8 (80)	9 (90)	7 (70)*	8 (80)		
	NBD	19 (95)	20 (100)	12 (60)*	15 (75)	17 (85)	20 (100)	10 (50)*	19 (95)		
Argument	SLP	10 (100)	7 (70)	10 (100)	9 (90)	10 (100)	9 (90)	8 (80)	8 (80)	10 (100)	7 (70)*
	NBD	19 (95)	18 (90)	20 (100)	20 (100)	20 (100)	14 (70)	18 (90)	15 (75)	19 (95)	9 (45)*
Directions	SLP	10 (100)	7 (70)	10 (100)	10 (100)	10 (100)	7 (70)	10 (100)	9 (90)*	7 (70)*	8 (80)
	NBD	15 (75)	15 (75)	19 (95)	19 (95)	18 (90)	17 (85)	18 (90)	10 (50)*	3 (15)*	16 (80)
Dishes	SLP	9 (90)	9 (90)	10 (100)	9 (90)	7 (70)*	9 (90)	7 (70)*			
	NBD	19 (95)	19 (95)	18 (90)	15 (75)	10 (50)*	15 (75)	9 (45)*			
Letter	SLP	10 (100)*	9 (90)*	10 (100)	9 (90)	9 (90)	10 (100)	10 (100)			
	NBD	11 (55)*	8 (40)*	20 (100)	16 (80)	16 (80)	16 (80)	17 (85)			

Note. BDAE = "cookie theft" picture from the Boston Diagnostic Aphasia Examination. WAB = "picnic" picture from the Western Aphasia Battery. Birthday, Cat = single pictures drawn for this project. Argument, Directions = picture sequences drawn for this project. Dishes, Letter = descriptions of procedures.
* = concepts that were eliminated from the final list because they were not mentioned by at least 70% of subjects in both groups.

thologist who had not been involved in the validation portion of the study. She was given the written scoring rules, the main concept lists and scoring examples, and several practice transcripts to score. After discussion of her scored practice transcripts, only minimal clarification was provided to this scorer as she scored the 24 transcripts selected for the first interjudge reliability assessment. Reliability values for this scorer equaled or exceeded those reported for the first scorer.

To assess intrajudge reliability, the first author rescored the transcripts of the same 12 aphasic subjects she had previously scored for the interjudge reliability assessment. Intrajudge reliability for each of the four scoring categories exceeded 80% (M = 90%). Intrajudge reliability exceeded 84% for each of the 12 aphasic subjects (M = 91%).

Test-retest Stability of Scores

The test-retest stability of main concept scores was evaluated in two ways. To provide a robust estimate of the effects of practice with the stimulus materials on subjects' main concept scores, the change in performance between two sessions that took place on the same day (Session 1 and Session 2) was measured. The changes in performance between Sessions 2 and 3 and between Sessions 1 and 3 also were measured. These latter differences provided an estimate of test-retest stability over a time interval that more closely resembles typical clinical practice, in which a patient's performance may be reassessed with the same materials after a week or more.

Absolute difference scores were calculated for each subject for the four scoring categories and also for a combination of the *Accurate/incomplete* and *Inaccurate* categories (AI + IN). (The latter combination represents all concepts that were mentioned, but were deficient in some way.) Absolute difference scores, rather than signed scores, were used so that negative differences would not cancel out positive differences when group statistics were calculated. Furthermore, the amount of change seemed more important to the issue of stability than its direction. The mean absolute difference scores for the 20 non-brain-damaged and 20 aphasic subjects are presented in Table 3. Difference scores generally were smallest for the Session 2 to Session 3 comparison and largest for the Session 1 to Session 2 comparison (Table 3).

To estimate the extent to which aphasic subjects' scores in one session were related to their scores in a subsequent session, Pearson product-moment correlation coefficients were calculated (Table 3). For the aphasic subjects, correlations ranged from .71 to .96. They generally were highest for the Session 2 to Session 3 comparison, and for the *Accurate/complete* and *Absent* scoring categories.

Although the correlation coefficients shown in Table 3 suggest a strong relationship among scores across the three sessions for the aphasic subjects, they do not indicate how accurately one could predict an individual's score in one session from his or her score in a different session. To provide this information, we calculated the standard error of measurement (SEM) for each of the four scoring categories and the AI + IN combined category using the formula:

$$SEM = SD\sqrt{1 - r}$$

where SD is the standard deviation for the distribution of obtained scores and r is the correlation coefficient. The SEM permits one to estimate the consistency (or reliability) with which a test will measure performance on repeated test occasions. The chances are about 68 in 100 that an individual's obtained score will not differ from the true score by more than ± 1 SEM, and about 95 in 100 that an individual's obtained score will not differ from the true score by more than ± 2 SEM. In general, the smaller the SEM, the greater the session-to-session stability of a score. The SEMs for the scoring categories evaluated in this study are provided in

TABLE 3. Mean absolute differences in scores across three sessions for non-brain-damaged and aphasic subjects and Pearson product-moment correlation coefficients (r) and standard error of measurement (SEM) values for aphasic subjects' scores.

		AB	AC	IN	AI	AI+IN
Non-brain-damaged (n = 20)						
Sessions 1 to 2	M	2.3	3.9	.7	2.7	2.6
	SD	1.8	2.7	.6	2.0	2.2
	Range	0–6	0–10	0–2	0–8	0–8
Sessions 1 to 3	M	1.8	2.9	.8	2.0	2.1
	SD	1.6	2.3	.6	1.7	1.6
	Range	0–6	0–10	0–2	0–5	0–5
Sessions 2 to 3	M	.4	1.7	.5	1.6	1.7
	SD	1.7	1.4	.5	1.6	1.6
	Range	0–3	0–5	0–1	0–5	0–5
Aphasic (n = 20)						
Sessions 1 to 2	M	2.7	3.6	1.9	3.3	4.1
	SD	1.7	2.5	1.8	1.8	2.7
	Range	0–7	0–11	0–6	0–6	0–10
	r	.94	.94	.77	.71	.72
	SEM	2	3	2	2	3
Sessions 1 to 3	M	2.7	2.9	2.2	2.5	3.2
	SD	1.8	2.6	1.6	1.6	2.2
	Range	0–7	0–12	0–6	0–5	0–7
	r	.94	.96	.74	.78	.71
	SEM	2	3	2	2	2
Sessions 2 to 3	M	2.1	3.3	1.8	2.5	2.6
	SD	1.8	2.3	1.3	2.3	1.7
	Range	0–6	0–7	0–5	0–8	1–6
	r	.96	.96	.82	.78	.86
	SEM	2	2	2	2	2

Note. AB = absent. AC = accurate/complete. AI = accurate/incomplete. NI = inaccurate.

Table 3. These SEM values suggest that one can assume with 68% confidence that an aphasic speaker's true score will lie within 2 points of the obtained scores for three of the scoring categories (AB, IN, and AI) and within 2 or 3 points of the obtained score for the other two scoring categories (AC and IN + AI). Doubling the SEM values in Table 3 yields the upper and lower limits of 95% confidence intervals for predicting an individual's true score from his or her expected score.

The range of scores across the non-brain-damaged subjects generally was very restricted and correlational analyses on these data were not appropriate. Table 4 gives the score ranges for the non-brain-damaged subjects.

Effects of Aphasia on Performance

To determine if the presence, completeness, and/or accuracy of main concepts distinguished the connected speech of aphasic speakers from that of non-brain-damaged speakers, the number of main concepts that received each of the four scores and the AI + IN combined score was calculated for the 20 aphasic subjects and the 20 non-brain-damaged subjects (Table 4).

As a group, aphasic subjects produced connected speech that had more *Absent* (AB), *Inaccurate* (IN), and *Accurate/incomplete* (AI) main concepts and fewer *Accurate/complete* (AC) main concepts in all three sessions than the non-brain-damaged group. To determine which of the Session 1 differences between the groups were statistically significant, a one-way analysis of variance, with groups as the main effect, was calculated within each of four individual score categories (AC, AI, IN, AB) and the combined AI + IN category. The familywise error rate was adjusted for multiple comparisons by setting the Type 1 (alpha) error rate for each test of significance at $p < .01$ (.05/5). The non-brain-damaged group received significantly more AC scores ($F = 40.47$), and significantly fewer AB ($F = 10.55$), IN ($F = 19.13$), AI ($F = 33.14$), and IN + AI ($F = 57.22$) scores than the aphasic group (degrees of freedom = 1,38 and $p < .01$ for all comparisons). The mean differences between the groups found in Sessions 2 and 3 either equalled or exceeded those found in Session 1. To avoid inflating the experimentwise Type 1 error rate, no statistical analyses of the Session 2 and 3 between-group differences were carried out.

Table 4 shows the number of aphasic subjects whose scores fell into the non-brain-damaged subjects' range for each scoring category. Scores for some aphasic subjects fell within the non-brain-damaged subjects' range for each scoring category. The number of aphasic subjects with scores in the non-brain-damaged subjects' score range was greatest for the *Absent* category, with over half of the aphasic

TABLE 4. Mean number (and percent) of main concepts (n = 53) that received each of four scores and one combination score for non-brain-damaged (NBD) and aphasic (APH) subjects in three sessions and the number of aphasic subjects' scores that fell in the non-brain-damaged subjects' range.

		AB	AC	IN	AI	AI+IN
Non-brain-damaged (n = 20)						
Session 1	M	7.4 (14)	39.7 (75)	1.1 (2)	4.9 (9)	6.0 (11)
	SD	4.3	7.6	1.1	3.9	4.2
	Range	0–15	27–51	0–3	0–14	0–14
	#APH in NBD Range	13	8	8	15	7
Session 2	M	5.7 (11)	42.3 (80)	.8 (2)	4.2 (8)	5.1 (10)
	SD	4.3	6.2	.7	2.6	2.6
	Range	0–15	31–51	0–2	1–10	2–11
	#APH in NBD Range	12	6	5	7	3
Session 3	M	6.1 (12)	41.8 (79)	.6 (1)	4.6 (9)	5.2 (10)
	SD	4.1	7.1	.7	3.0	3.3
	Range	1–13	30–50	0–3	1–10	1–11
	#APH in NBD Range	11	7	5	8	3
Aphasic (n = 20)						
Session 1	M	14.6 (28)	21.0 (40)	5.2 (10)	12.3 (23)	17.5 (33)
	SD	8.6	10.3	4.0	4.1	5.2
	Range	3–33	3–36	1–17	2–20	9–26
Session 2	M	14.2 (27)	20.5 (39)	5.1 (10)	13.3 (25)	18.4 (35)
	SD	9.1	12.4	3.8	5.1	5.8
	Range	0–30	2–42	0–16	4–22	7–27
Session 3	M	13.3 (25)	22.5 (42)	5.3 (10)	12.3 (23)	17.6 (33)
	SD	8.0	11.6	2.8	4.5	4.9
	Range	3–31	4–48	0–13	4–20	8–26

Note. AB = absent. AC = accurate/complete. AI = accurate/incomplete. IN = inaccurate.

subjects' *Absent* scores falling in the non-brain-damaged subjects' range in all three sessions. The number of aphasic subjects with scores in the non-brain-damaged subjects' score range was smallest for the combined IN + AI category in all three sessions.

More of the aphasic subjects' scores fell within the non-brain-damaged subjects' range in Session 1 than in Sessions 2 or 3. This appears to be accounted for primarily by the non-brain-damaged subjects' small but generally positive changes in production of main concepts following their first experience with the elicitation stimuli. The non-brain-damaged groups' production of *Accurate/complete* main concepts increased from Session 1 to Sessions 2 and 3, with a consequent decrease in the number of *Accurate/incomplete*, *Inaccurate*, or *Absent* main concepts. Although the aphasic group showed a slight decrease in the number of *Absent* main concepts from Session 1 to Sessions 2 and 3, they did not exhibit consistent changes in the other scoring categories.

To evaluate the strength of the relationship between aphasic subjects' estimated aphasia severity and their main concept performance, Pearson correlation coefficients were calculated between subjects' SPICA overall percentiles and the number of *Absent*, *Accurate/complete*, or *Accurate/incomplete* plus *Inaccurate* (AI + IN) main concepts in Session 1. The correlation between SPICA overall percentile and the number of *Absent* main concepts yielded $r = -.91$ and the correlation between SPICA overall percentile and the number of *Accurate/complete* main concepts yielded $r = .82$, suggesting that estimated severity of aphasia is strongly related to the number of both *Absent* and *Accurate/complete* main concepts. The correlation between SPICA overall percentile and the number of *Accurate/incomplete* plus *Inaccurate* (AI + IN) main concepts yielded $r = .12$, suggesting that estimated severity of aphasia had no important relation to the number of *Accurate/incomplete* plus *Inaccurate* main concepts produced.

Discussion

The results of this study suggest that the presence, completeness, and accuracy of main concepts in connected speech can be scored with acceptable interjudge and intrajudge reliability for most adults with aphasia. The relatively low interjudge agreement (74%) for scoring the most impaired aphasic subject's speech samples suggests that main concept scoring may be less reliable for adults with moderately-severe to severe aphasia. The scoring categories of absent and accurate/incomplete were involved in the majority of the disagreements for this subject. This supports the scorers' subjective feeling that the empty, paraphasic nature of this speaker's copious output made it difficult to determine if he was attempting a particular main concept.

Main concept scores appear to be reasonably stable from session to session for most adults with aphasia, at least when they are based on responses to 8 stimuli such as those described herein. In previous studies (Brookshire & Nicholas,

1994a, 1994b), we found acceptable test-retest stability for the percent correct information units measure for most adults with aphasia when speech sample size was reduced by decreasing the number of speech elicitation stimuli from 10 to 4 or 5. At this time we have not evaluated whether test-retest stability for main concept scores will remain acceptable if fewer stimuli are used to elicit the speech samples from which the main concept measures are obtained, although it seems likely that substantially reducing the number of elicitation stimuli may compromise test-retest stability. Regardless of the number of stimuli used to elicit speech samples from which main concept scores are derived, those who wish to assess changes in an individual aphasic adult's main concept production over time should establish the test-retest stability of main concept scores for that individual across several baseline sessions, because some individual subjects may exhibit relatively large changes from session to session.

Results of previous studies have shown that groups of brain-damaged subjects produced fewer main units of information in narrative discourse than non-brain-damaged subjects. Our results showed a similar pattern, at least at the level of group means (Table 4). However, when the performance of individual aphasic subjects was evaluated, most of their scores for number of *Absent* main concepts (from 55% to 65%) fell within the range for the non-brain-damaged subjects. There generally was much less overlap between the groups for scores in the other categories. The least overlap was obtained when the number of *Accurate/incomplete* main concepts was added to the number of *Inaccurate* main concepts (AI + IN). These results show that it was not the number of main concepts that subjects failed to mention that best distinguished aphasic subjects' performance from that of their non-brain-damaged counterparts, but rather the completeness and accuracy of the main concepts they did produce.

In previous studies, descriptions of scoring procedures did not specify the levels of completeness and accuracy required for main units of information to be scored as present in the speech samples. It is possible that in those studies, incomplete units of information or units of information containing inaccuracies may have been scored as absent, leading to the imprecise conclusion that aphasic adults leave out main units of information more often than non-brain-damaged adults.

Although the least overlap in performance between aphasic and non-brain-damaged speakers occurred for the combined AI + IN scoring categories, the number of main concepts receiving these scores in the transcripts of speakers with aphasia was not strongly correlated with aphasia severity. Instead, the number of main concepts scored either as *Absent* or as *Accurate/complete* was much more strongly related to aphasia severity, with speakers whose aphasia was less severe generally having fewer *Absent* main concepts and a larger number of *Accurate/complete* main concepts than those with more severe aphasia. The absence of a strong relationship between inaccurate and/or incomplete main concepts and aphasia severity may relate to the fact that main concepts within each of these categories actually represent a continuum of accuracy or completeness. Some concepts might be inaccurate because they contain a single inaccurate essential element, whereas others might be inaccurate because they contain a numbers of inaccurate essential elements. Likewise, some concepts might be incomplete because the speaker left out a single essential element, whereas others might be incomplete because several essential elements were missing. It seems likely that if one were to rank main concepts falling into these two categories on some scale of "goodness" a stronger relationship between them and aphasia severity would be found.

Nevertheless, the results of this study suggest that most adults with aphasia are sensitive to macrostructural aspects of discourse, because they usually attempted to communicate most of the main concepts called for by a topic or elicitation stimulus, even though their production of the concepts may be incomplete or contain inaccuracies. However, as the severity of an individual's aphasia increases, the number of missing main concepts tends to increase. The three aphasic subjects who attempted fewer than half of the main concepts had the lowest SPICA percentiles in the group (40th to 41st percentile). This may suggest that, even though adults with aphasia are sensitive to the need for producing main concepts in their connected speech, formulation and production problems associated with increasing severity of aphasia may eventually compromise their ability to produce them.

Measures of a speaker's main concept production can provide a useful complement to other measures of communicative informativeness and efficiency, such as percent correct information units or content units per minute. These latter measures provide information about how much of what a speaker says is accurate, relevant, and informative, but they do not provide information about how well the speaker communicates what is most salient about an elicitation stimulus or topic. A speaker may be very efficient, in that almost all the words he or she produces are accurate, relevant, and informative, but the speaker may fail to convey the main units of information that give connected speech its overall structure, or "point." On the other hand, a speaker may provide most or all of the main units of information, but do so inefficiently, because of false starts, unnecessary repetition, irrelevant comments, or filler words and phrases. When the presence, accuracy, and completeness of main concept production are considered together with the efficiency with which a speaker conveys information, a more complete picture of aphasic adults' strengths and weaknesses in connected speech production can be obtained. Such a combination of measures may be useful in focusing treatment and in measuring changes in the informativeness and efficiency of connected speech as a consequence of treatment.

Other measures, such as subjective ratings of the adequacy of local and global coherence and objective measurement of various types of "performance deviations" (Loban, 1976), may add other relevant dimensions to the picture. It seems likely that listeners might be able to tolerate higher levels of performance deviations, such as false starts, unnecessary repetition, or inaccurate or nonspecific words, if the main information in a message has been presented in a coherent manner.

Research on the relationship among measures of main concept production, performance deviations, and listeners'

subjective judgments of speakers' communicative effectiveness might provide a theoretical basis for future treatment studies. If main concept production is found to be a strong determinant of listeners' judgments of speakers' communicative effectiveness, then aphasia treatment studies might focus on evaluating the effects of training adults with aphasia to identify the most salient information in their messages and to emphasize its production over that of other, more peripheral information.

Acknowledgments

This research was supported by the Department of Veterans Affairs Medical Research Service. We thank Shelley Brundage, Julia Edgar, Donald MacLennan, and Annalisa Margheri for their assistance with this study.

References

Brookshire, R. H., & Nicholas, L. E. (1994a). Test-retest stability of measures of connected speech in aphasia. *Clinical Aphasiology, 22*, 119–133.

Brookshire, R. H., & Nicholas, L. E. (1994b). Speech sample size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research, 37*, 399–407.

Disimoni, F. G., Keith, R. L., & Darley, F. L. (1980). Prediction of PICA overall score by short versions of the test. *Journal of Speech and Hearing Research, 23*, 511–516.

Duffy, J. R., Keith, R. L., Shane, H., & Podraza, B. L. (1976). Performance of normal (non-brain-damaged) adults on the Porch Index of Communicative Ability. In R. H. Brookshire (Ed.), *Clinical Aphasiology: Conference Proceedings* (pp. 32–42). Minneapolis: BRK Publishers.

Gleason, J. B., Goodglass, H., Obler, L., Green, E., Hyde, M. R., & Weintraub, S. (1980). Narrative strategies of aphasic and normal-speaking subjects. *Journal of Speech and Hearing Research, 23*, 370–382.

Goodglass, H., & Kaplan, E. (1983). *The Boston Diagnostic Aphasia Examination.* Boston: Lea & Febiger.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English.* London: Longman.

Heir, D. B., Hagenlocker, K., & Shindler, A. G. (1985). Language disintegration in dementia: Effects of etiology and severity. *Brain and Language, 25*, 117–133.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels: Research report #3.* Champaign, IL: National Council of Teachers of English.

Joanette, Y., Goulet, P., Ska, B., & Nespoulous, J. (1986). Information content of narrative discourse in right-brain-damaged right-handers. *Brain and Language, 29*, 81–105.

Johnson, R. E. (1970). Recall of prose as a function of the structural importance of linguistic units. *Journal of Verbal Learning and Verbal Behavior, 9*, 12–20.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test.* Philadelphia: Lea & Febiger.

Kertesz, A. (1982). *The Western Aphasia Battery.* New York: Grune & Stratton.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review, 85*, 363–394.

Loban, W. (1976). *Language development: Kindergarten through grade twelve.* Urbana, IL: National Council of Teachers of English.

Meyer, B. J. F. (1975). *The organization of prose and its effects on memory.* Amsterdam: North-Holland.

Meyer, B. J. F., & McConkie (1973). What is recalled after hearing a passage? *Journal of Educational Psychology, 65*, 109–117.

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research, 36*, 338–350.

Nicholas, L. E., Brookshire, R. H., MacLennan, D. L., Schumacher, J. S., & Porrazzo, S. A. (1989). Revised administration and scoring procedures for the *Boston Naming Test* and norms for non-brain-damaged adults. *Aphasiology, 3*, 569–580.

Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech and Hearing Research, 28*, 405–410.

Porch, B. E. (1971). *The Porch Index of Communicative Ability.* Palo Alto, CA: Consulting Psychologists Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, goals, plans, and understanding.* Hillsdale, NJ: L. E. A. Publishers.

Schuell, H. (1972). *The Minnesota Test for Differential Diagnosis of Aphasia.* Minneapolis: University of Minnesota Press.

Shewan, C. M. (1988). The Shewan Spontaneous Language Analysis (SSLA) System for aphasic adults: Description, reliability, and validity. *Journal of Communication Disorders, 21*, 103–138.

Stein, J., & Urdang, L. (Eds.). (1973). *The Random House dictionary of the English language.* New York: Random House.

Ulatowska, H. K., Freedman-Stern, R., Doyel, A. W., & Macaluso-Haynes, S. (1983). Production of narrative discourse in aphasia. *Brain and Language, 19*, 317–334.

van Dijk, T. A. (1980). *Macrostructures.* Hillsdale, NJ: Lawrence Erlbaum.

van Dijk, T. A. (1985). Semantic discourse analysis. In T. A. van Dijk (Ed.), *Handbook of discourse analysis* (Vol. 2; pp. 103–136). New York: Academic Press.

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension.* New York: Academic Press.

Wagenaar, E., Snow, C., & Prins, R. (1975). Spontaneous speech of aphasic patients: A psycholinguistic analysis. *Brain and Language, 2*, 281–303.

Yorkston, K., & Beukelman, D. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of Speech and Hearing Disorders, 45*, 27–36.

Received October 13, 1993
Accepted July 1, 1994

Contact author: Linda E. Nicholas, MA, Aphasia Research (127A), VA Medical Center, One Veterans Drive, Minneapolis, MN 55417. (Copies of the eight main concept lists and sets of scoring examples can be obtained from the contact author.)

Appendix A

Rules for Scoring Main Concepts

You will be deciding if the essential information in each main concept for a story or procedure was mentioned in an accurate and complete form. For each main concept you will give one of four scores, depending on the presence, accuracy, and completeness of the essential information in the main concept.

AC: accurate and complete
AI: accurate but incomplete.
IN: inaccurate.
AB: absent.

The main concepts for the 8 elicitation stimuli are provided in a set of main concept lists. In the main concept lists, the most common alternate words or information are shown in parentheses. Essential information for each concept is underlined. Examples of statements that would receive AC, AI, or IN scores are provided for each main concept in example lists.

Main concepts do not have to be given in the transcript in the same order they are shown in the main concept list.
If a main concept is spoken several times, score the final version.

Definitions of Scores

AC: Accurate, complete
All essential information is accurate and complete.

AI: Accurate, incomplete
Part of the essential information is accurate, but one or more essential parts are missing.

IN: Inaccurate
One or more parts of the essential information are inaccurate. If a main concept that contains inaccurate essential information also has missing essential information, you can make a note of this (e.g., IN-IC), but it will not affect the score. (We found that scoring incompleteness when essential information is inaccurate is difficult to do reliably.)

AB: Absent
None of the essential information is given. The speaker has said nothing that appears to be an attempt to communicate the essential information in the main concept.

Rules for Determining Accuracy and Completeness

Accuracy

Wording. The wording of essential information does not have to be the same as that of the listed main concept, but the general meaning must be the same. For example, for the main concept, "The woman is doing the dishes," the following statements would be considered accurate:

—The mother is cleaning the dishes.
—The maid is drying a plate.
—The lady was washing her dishes.

These statements are all plausible, given the stimulus picture. Some alternate versions of essential information in specific main concepts are given in parentheses in the main concept lists. Figures of speech are acceptable.

Grammatical form and word order. Essential information does not have to be spoken in standard grammatical form or standard word order, as long as deviations would not lead to miscomprehension of the essential meaning of the main concept. For example, the following statements would be considered accurate:

—Mother do dishes.
—Lady drying some plate.
—Woman dishes drying them.

Articulation. Essential words do not have to be correctly articulated to be considered accurate, as long as they would be intelligible to a listener as the target words in the context of what the speaker is saying. Assume that the listener has seen the stimulus picture or knows which procedure is being discussed.

Inaccurate words in accurate main concepts. If the essential information in a main concept is accurate, but inaccurate words also are included in the concept, consider the main concept accurate unless the inaccurate words alter essential information to make it inaccurate. For example, if someone said, "The woman is drying the dishes with a bath towel," the concept would be considered accurate even though the woman is using a dish towel, because the essential information is accurate.

Effects of statement form on essential information. The form in which a speaker chooses to state a main concept affects which information is considered essential. For example, someone describing how to do dishes can say the first main concept as "Put water in the sink," or "Run the water." If the speaker said, "Put water in the kitchen box," the main concept would be considered inaccurate, because the verb "put" requires a preposition and an object (which must be correct). However, if the speaker said, "Run the water in the kitchen box," the main concept would be considered accurate, because the verb "run" does not require a preposition and an object.

Inaccurate pronoun referents. If a referent is inaccurate but a pronoun that refers to it is accurate, consider the statement with the pronoun accurate. For example, if someone said, "The man is washing dishes. She is not paying attention to her kids," the first statement would be considered inaccurate but the second would be considered accurate.

Restatement of essential information. If essential information is stated one way and then restated, score the final version spoken. For example, if someone said, "The woman he is doing the dishes," the statement would be considered inaccurate. If the speaker said, "The man no woman is doing the dishes," the statement would be considered accurate.

Completeness

Missing essential information. If all of the essential information for a main concept is not mentioned, either in its list form or in a form that has the same general meaning, consider the main concept incomplete.

Statements containing some of the essential information. If a statement that is not listed as a main concept contains some of the essential information for a main concept, consider the main concept incomplete. For example, if someone said, "The woman is standing there," "woman" would be considered an incomplete main concept for "The woman is doing dishes." This would be the case as long as "woman" was not counted toward another main concept and the character has not been mentioned previously. Such statements serve to establish a character, element, or action in a story.

Nonspecific words. If nonspecific words are spoken in place of essential information, consider the main concept incomplete. For example, "The woman is washing that thing," would be considered an incomplete main concept. In some cases there may be legitimate uncertainty about a specific person or element in a picture (e.g.,

"Someone is fishing" (WAB). In such instances, the nonspecific word will appear on the main concept list and its use will not render the main concept incomplete.

Pronoun referents. If the referent for a pronoun is ambiguous, consider the main concept that contains the ambiguous pronoun incomplete, but only for the first use of the pronoun for that referent. Do not consider a main concept incomplete if there is no antecedent referent for a pronoun but the referent is clear from context. For example, if there is only one "she" or "he" or "they" in a picture, the pronoun would be unambiguous without an antecedent referent. (This is different from the Correct Information Unit rule that requires the activation of a referent prior to first use of a pronoun in place of that referent, regardless of the picture context.)

Procedure for Scoring Main Concepts

1. Familiarize yourself with the stimulus pictures, the main concept lists, and the main concept scoring examples list.

2. Begin scoring a transcript by bracketing the information for each main concept. Then go back and put the appropriate main concept number and the score for each main concept above the bracketed information, as in the following examples:

1 AC

". . . it looks like [the woman is doing the dishes] . . ."

2 AC

". . . [the water is running onto the floor] by her feet . . ."

3. Write the numbers for those main concepts that are absent in the right margin just below the last line for each transcript.

Appendix B

Main Concepts and Scoring Examples for the BDAE Cookie Theft Picture

1. THE WOMAN (MOTHER) IS DOING DISHES.

AC
Nancy washing the dishes.
The mother is drying a plate.
The maid is cleaning the dishes.
Mother do dishes.

AI
The mother is standing by the sink.
Mother dishes.
Lady with dishes.

IN
The woman is washing clothes.
The mother he is washing dishes.

2. THE SINK (WATER) IS OVERFLOWING (RUNNING OVER).

AC
The woman let the water run over.
Water is spilling out.
The water is running on the floor.
The man left the water on and it's running on the floor. ("Man" is inaccurate but nonessential.)

AI
Some problem with the water.
Left the water running.
Water on the floor.

IN
The water is jumping over.
The water is running on the ceiling.
The nadercom is running over. (You should feel confident that an unintelligible word is an attempt at the target word before giving an IN score.)

3. THE BOY IS ON A STOOL.

AC
Boy on stool.
The stool is tipping over with the boy. (#5 AC also)
Johnny is going to fall from the stool. (#5 AC also)
He is on the high stool. (Score as AC even if no referent has been established for "he" because there is only one male character in the picture.)

AI
One of them is on a stool.
A stool there.
A boy is on that thing.

IN
The boy is on a chair.
The girl is on a stool.
They are on the stool.

4. THE BOY (KIDS) IS GETTING (STEALING) COOKIES (GETTING INTO THE COOKIE JAR).

AC
Junior is climbing where he shouldn't be to get cookies.
Kids getting into cookie jar.
Jack is up in the cookie jar.

AI
Her daughter and son want to have cookies.
Junior is up in the cupboard.
Cookie jar cabinet.
Two kids cookie jar.

IN
A boy about to steal cookie jars.
The children are reaching into the cookie jar.
They are climbing up to get some cookies.

5. THE STOOL IS TIPPING. (THE BOY IS FALLING.)

AC
The stool is slipping to the side.
The stool is about to tip over.
The kid is about to land on his backside.
Stool tipping.
His stool is tipping. (No credit for #3).

AI
That thing there is tipping.
One of them is falling.
The girl is gonna let him fall.

IN
He spilled the footer.
They almost fell off the stool.
These guys will fall over.
Boy fell.

6. THE GIRL IS REACHING FOR A COOKIE. (The BOY HANDS THE GIRL A COOKIE.) *OR SOME MENTION OF A PLAUSIBLE ACTION BY THE GIRL OR LOCATION OF THE GIRL

AC
Boy getting a cookie for sister.
*The girl has her hand out.
*Little girl in the kitchen too.
*Sally is telling him "shh."

AI
The little girl should be holding the stool but she isn't.
Boy give girl.
Two children—one is on the floor.
Sister wants a cookie.

IN
She was handing the cookie to the girl.

7. THE WOMAN (MOTHER) IS NOT NOTICING (PAYING ATTENTION).

AC
The mother appears to be lost in her own thoughts.
This housewife seems to be daydreaming.
The father—no mom is not paying attention (Self correction)

AI
Mother has her back turned.
The woman is looking out the window.
I see an apparently unfrustrated mother standing in the middle of chaos.

IN
The mother is angry.
The father is daydreaming.

Note. The numbered statements are the main concepts. The essential information for each main concept is underlined. In the IN sections of the example list, underlined words are those that are inaccurate.