

Relevance in the Language Production of Aphasic Patients

Julie Ann Christiansen

Klinik Burg Landshut

Recent studies have shown that even mildly impaired aphasic patients of all types produce distinctive patterns of coherence violations in discourse. Using three tests tapping the ability of Broca's, conduction, and Wernicke's aphasics to maintain coherence in production at the word, sentence, and discourse level, the current study questions whether the intrusion of irrelevant utterances might result from a general processing impairment affecting both local coherence between consecutive sentences and word associations in the lexicon, or whether it is related only to discourse structure. Results indicate that the coherence problems of the Wernicke's aphasics are probably limited to discourse, although more severely impaired patients may experience difficulty at every level of language production. Furthermore, Broca's aphasics produced a ratio of essential and peripheral propositions roughly equivalent to the normal control subjects, indicating that they do not usually confine themselves to producing only the most crucial information to a given topic, even though the information content was greatly reduced.

Relevancia en la Producción del Lenguaje de Pacientes Afásicos

Recientes estudios han mostrado que los diferentes tipos de pacientes afásicos leves producen diferentes patrones de violaciones en la coherencia del discurso. Este estudio cuestiona si la intrusión de alteraciones irrelevantes puede resultar de un procesamiento general afectado que involucre tanto la coherencia local entre frases consecutivas y asociaciones entre palabras en el lexicon, o si está relacionado únicamente con la estructura del discurso. Los resultados indican que los problemas de coherencia de los afásicos de Wernicke están probablemente limitados al discurso, aunque los pacientes con daños severos pueden experimentar dificultades en cada nivel de producción del lenguaje. Adicionalmente, los afásicos de Broca producen una tasa de proposiciones esenciales y periféricas equivalentes a los sujetos controles, indicando que ellos no se

Correspondence to the first author should be addressed to Abtlg. Sprachtherapie, Klinik Burg Landshut, Kueser Plateau, D-54463 Bernkastel-Kues, Germany. Email: julie.christiansen@t-online.de

The author wishes to acknowledge Sui Choi for drawing the cartoon stimuli and Lorrie Verplaetse, Pam Wendler-Shaw, Todd Brown, Susan Ward, Qian Hu, and Sulochana Naidoo for scoring the data. The author would also like to thank Harold Goodglass and Herman Kolk for their helpful comments and support.

limitan usualmente a producir únicamente la información más importante sobre un tema, aunque el contenido de la información se redujo.

Coherence can be defined as the semantic connectedness of a text (Van Dijk, 1977; 1980). For a story narrative to be coherent, it should contain all of the propositions necessary for a listener to construct the correct macrostructure for the story, but should not include irrelevant and tangential propositions which would distract the listener from the main points (see Christiansen, 1995; Kintsch and Van Dijk, 1978; Van Dijk and Kintsch, 1983). Coherence has been thought to be relatively well-preserved in the discourse of aphasic patients (Glosser and Deser, 1990; Huber, 1990; Ulatowska, Allard, and Chapman, 1990; Ulatowska, Freedman-Stern, Doyel, Macaluso-Haynes, and North, 1983; Ulatowska, North, and Macaluso-Haynes, 1981), however, this may not pertain to all types of aphasic syndromes. Broca's aphasics, while producing few irrelevant propositions, have difficulty producing all of the propositions essential to a story (Goodglass, Christiansen, and Gallagher, 1994).

In a propositional analysis of the production of fluent aphasic patients, Christiansen (1995) found that particularly Wernicke's aphasics produced numerous irrelevant propositions in their narratives. Christiansen proposed that the Wernicke's aphasics had difficulty determining which propositions were crucial to a particular plot. Because of this uncertainty, these patients may simply have produced every proposition which came to mind, essential or irrelevant, as an adaptive strategy. This strategy ensured that their narratives were complete, at the expense of disrupting coherence.

It is not clear whether the irrelevant propositions produced by Wernicke's aphasics represent a particular disturbance in discourse coherence or is secondary to an overall semantic impairment affecting language production at the level of words and sentences as well as paragraphs. By using a multiple case study design, the current study seeks to determine the relations between the ability of individual aphasic patients to maintain topic-relevant language output in narrative discourse, consecutive sentences, and word associations. Specifically, the following questions are asked:

1. Can Broca's, conduction, and Wernicke's aphasics produce the relevant story information in narratives, without the intrusion of irrelevant propositions, despite their difficulties in formulating grammatical sentences and accessing appropriate lexical items?
2. Are difficulties in producing coherent discourse associated with particular aphasic syndromes?

3. If an aphasic patient produces an abnormal number of irrelevant propositions in discourse, is this intrusion of irrelevant propositions confined to the production of narrative-length discourse, or might it result from a more general processing impairment also affecting the production of smaller units such as single words or sentences?

The Function of Relevance in Discourse

Relevance has been shown to be a key tool in all realms of human communication, including action sequences and verbal comprehension, as well as verbal production (Sperber and Wilson, 1986). But how does a speaker know what is relevant to the current discourse, and how can a listener use relevance to interpret the speaker's message? In Sperber and Wilson's (1986) relevance theory, each participant in a conversation begins with his/her cognitive environment, which is the set of all assumptions available to the individual at any particular time. These assumptions include facts about the current physical environment, encyclopedic facts, common experiences with the other participants, etc., as well as all inferences made from this knowledge. When an individual encounters new information, that information, along with all of the inferences made from it, causes a change in the individual's cognitive environment, which Sperber and Wilson call *contextual effects*. When a proposition is relevant, as defined by Sperber and Wilson, it produces contextual effects on the listener's cognitive environment with a minimum of processing effort--the greater the effects, the greater the relevance.

To identify the main points, or macrostructures, of a given piece of discourse, the listener must first be able to parse the segment into a series of semantic propositions (Van Dijk and Kintsch, 1983), defined as the minimal semantic units represented therein. As specified in Kintsch (1991), propositions consist of a head and a number of slots for arguments, along with their precise relation to the head. Semantic connections are established between propositions by means such as bridging (Clark, 1975; Clark and Haviland, 1977; Haviland and Clark, 1974), argument overlap (Kintsch and Van Dijk, 1978), and centering (Grosz and Sidner, 1986; Gordon, 1993), to name a few. Yet inherent throughout each of these means of extracting the main points of discourse is the presumption that the propositions presented are somehow relevant to the main points of the discourse, and that the listener is able to discern the degrees of relevance for each proposition, based on the presentation by the speaker and the current cognitive

environment of the listener.

As relevance and macrostructure theories have focused primarily on discourse comprehension, less is known about the processes involved in discourse production. Production processes cannot simply be the reverse of comprehension processes, as the formation of macrostructures in production is internally driven, while the extraction of macrostructures in comprehension must be deduced from external information in the discourse context. Yet the two processes must be closely related, as both are drawing upon the same rules of macro- to microstructure mapping. It seems reasonable to assume that in language production, the speaker begins by generating a global intent, or macroproposition (Levelt, 1989; Van Dijk and Kintsch, 1983). Once the global intent has been formed, the speaker must then specify the information for smaller, more local units (e.g., local utterances and micropropositions). Van Dijk and Kintsch (1983) have suggested three possible operations which the speaker might use: adding details, particularizing general propositions, and analyzing (i.e., breaking down complex actions into simpler components). These three operations are roughly the reverse of the three macrorules used in comprehension: deletion, generalization, and construction.

Relevance in the Discourse of Aphasic Patients

In past studies, coherence problems in discourse have usually been ascribed to stroke patients with right-hemisphere brain damage (RBD), who demonstrate striking coherence and pragmatic difficulties despite an intact ability to encode and decode language at the surface level (Myers, 1993; Weylman, Brownell, and Gardner, 1988). In contrast to the left-hemisphere brain-damaged (LBD) aphasic patients, RBD stroke patients often have difficulty interpreting connotative and metaphorical meanings in words and sentences (Brownell, Potter, Michelow, and Gardner, 1984; Brownell, Simpson, Bihrlé, Potter, and Gardner, 1990; Van Lancker and Kempler, 1987; Winner and Gardner, 1977), as well as comprehending indirect requests (Foldi, 1987; Hirst, LeDoux, and Stein, 1984; Weylman, Brownell, Roman, and Gardner, 1989), and jokes (Bihrlé, Brownell, Powelson, and Gardner, 1986; Brownell and Gardner, 1988; Brownell, Michel, Powelson, and Gardner, 1983). Furthermore, they have difficulty inferencing from connected discourse (Brownell, Potter, Bihrlé, and Gardner, 1983; Molloy, Brownell, and Gardner, 1990). However, the coherence difficulties demonstrated by RBD and LBD stroke patients appear to be more divergent in comprehension than in production. In production, both LBD aphasic and

RBD non-aphasic patients demonstrate reduced informational content in their connected discourse (Gleason, Goodglass, Obler, Green, Hyde, and Weintraub, 1980; Joannette, Goulet, Ska, and Nespoulous, 1986; Joannette and Goulet, 1990; Ulatowska, et al., 1981; Ulatowska, et al., 1983). In addition, both Wernicke's aphasics and RBD patients produce numerous extraneous utterances in their narratives (Christiansen, 1995; Myers, 1993). The quality of the extraneous utterances, however, seems to differ among the two groups. While RBD patients may relate a series of details, including tangential and irrelevant observations, without actually getting to the main point of the story, Wernicke's aphasics can relate the essential elements of the story even though the story is infiltrated with numerous irrelevant details (Christiansen, 1993; 1995). In a few cases, aphasic patients have been found to confabulate responses (Sandson, Albert, and Alexander, 1986). However, these confabulations may be due to other neuropsychological impairments, such as attention and memory deficits, not directly resulting from the given language deficits.

In analyzing the LBD aphasic patients' ability to incorporate Sperber and Wilson's principle of relevance into their narratives, it is necessary to consider both the lack of essential propositions and the intrusion of extraneous propositions. Most aphasic patients demonstrate some difficulty either in providing all of the essential propositions of a story or in carefully monitoring the relevance of additional propositions to the story being told, with the possible exception of very mild patients and some conduction aphasics (Christiansen, 1994; 1995).

Story Content and Information Gaps. When compared to age-matched normal control subjects, many aphasic patients fail to produce an equivalent number of essential propositions. In narratives elicited with cartoon pictures, both Broca's and Wernicke's aphasics consistently produce fewer main points than nonaphasic controls and tend to perseverate on the most salient points in the story (Gleason, et al., 1980). Even patients who include most of the essential propositions of the story often show a marked reduction in nonessential elaborative propositions as compared to the normal subjects (Ulatowska, et al., 1983). Earlier studies of coherence in aphasia tended neither to investigate individual behaviors of the patient population nor to analyze the patient's performance in relation to the subtypes of aphasia. In a propositional analysis of Broca's, amnesic, and Wernicke's aphasics, Huber (1990) found that the aphasics produced a similar range of essential and optional propositions. However, from the individual data he presents, one can see that the Broca's and anomic aphasics in his study also tended to omit

essential story propositions. Additional studies of propositional content in narrative production revealed information gaps in the narratives of both Broca's aphasics (Goodglass, et al., 1994) and anomic aphasics (Christiansen, 1995), confirming the tendency for some aphasic patients to omit essential propositions from their narratives.

Irrelevant Propositions. Though they rarely omit essential information in narrative production, Wernicke's aphasics, as opposed to other types of aphasics, seem to have the most difficulty maintaining relevance in their stories (Christiansen, 1995). The anomic, conduction, and Wernicke's aphasics in Christiansen's study all produced distinctively different discourse patterns. The anomic aphasics tended to give very brief narratives, often omitting essential propositions and refraining from unnecessary propositional elaborations. The conduction aphasics did not omit essential propositions in their narratives, but reiterated propositions with abnormal frequency. Christiansen attributed the narrative production patterns of the anomic and conduction aphasics to the strategies they used to compensate for their lexical and syntactic deficits. The Wernicke's aphasics, on the other hand, produced all of the essential propositions required for their narratives, but in addition, produced significantly more irrelevant propositions than either of the other aphasic groups or the normal controls. Christiansen concluded that the production of all possible narrative propositions, relevant and irrelevant, by the Wernicke's patients may also be a production strategy, but one that is used to compensate for a loss of macrostructural knowledge. Because the Wernicke's could no longer judge the relevance of each of their propositions to the main points of the story, they produced everything which came to mind. Such a production strategy would provide completeness at the expense of coherence.

In light of the fact that some aphasic patients exhibit difficulty in producing only those propositions that are most relevant to a particular story, the current study seeks to investigate whether those individual patients who include irrelevant propositions in their narratives also produce irrelevant responses when asked to produce only sentences or individual words in a particular context. Such behavior would indicate that the intrusion of irrelevant propositions into the patient's discourse might be related to deficits in interpreting the conceptual relevance of words and propositions to a particular topic. On the other hand, it is possible, that the intrusion of irrelevant propositions in narrative is confined to discourse, in which case the inability to maintain relevance may result from an overload in discourse processing when the subject is required to plan lengthy connected speech.

Methods

This current study presents three subtests from Christiansen's (1994) Relevance/ Coherence Battery (RCB). The RCB was originally designed to analyze the aphasics' ability to utilize commonly known scripts and frames to process language, with a primary emphasis on relevance and contextualization. The original battery consisted of nine subtests which focused on the subjects' ability to maintain relevance at three levels: 1) accessing situational frames to comprehend and produce single words, 2) accessing scripts and frames both to produce and judge the relevance of sequential utterances or events, and 3) interpreting and producing coherent narrative discourse. At each level, the RCB tested both verbal and nonverbal processing abilities. As this study focuses solely on relevance in verbal production and the issue of processing load, only the three production subtests from the RCB are used: narrative production, sentence pair completion, and word list generation.

Subjects

The subjects used in this study consisted of 13 mildly to moderately impaired aphasic patients (3 female, 10 male), as described in Table 1, and exhibited various types of aphasic syndromes (4 Broca's, 4 conduction, and 5 Wernicke's). Each subject was evaluated at the time of testing with the Boston Diagnostic Aphasia Exam (BDAE) (Goodglass and Kaplan, 1983), the short form of the Token Test (DeRenzi and Faglioni, 1978), and the Boston Naming Test (BNT) (Goodglass, Kaplan, and Weintraub, 1983). All of the subjects demonstrated relatively good auditory comprehension (averaging at least in the 70th percentile on the BDAE auditory comprehension subtests). All subjects had experienced either a single, left hemisphere infarction or a ruptured aneurysm, and none of the subjects had any other history of head injury, neurological disease, or active alcoholism.

Table 1.
Background Information for Aphasic Subjects

Sub	Sex	Age	Edu	Hand ^a	TPO	S.R	Com	Nam	Rep	BNT	TT	LESION SITE
MA	Mm	75	8	R (R)	16.7	1.5	78.3	85.3	83.3	40	57	Left posterior frontal lobe, patchy in parietal lobe
RO	M	64	10	L (R)	5.3	2.0	77.5	73.7	63.3	31	61	Left cortical and subcortical fronto-parietal region
BL	M	67	12	R (R)	15.8	2.0	87.5	81.7	35.0	40	69	Left Broca's area with subcortical extension
MP	F	72	13	R (L)	3.5	1.5	91.3	93.0	86.7	45	74	Left frontal border zone of anterior MCA
PA	F	71	12	R (R)	2.8	3.5	92.5	97.3	73.3	47	77	Left fronto-parietal area with PVWM extension
CA	M	60	10	R (R)	3.5	2.5	89.0	92.0	73.3	27	75	Left SM gyrus and most subcortical structures
GO	F	72	12	R (L)	2.7	2.5	90.0	76.7	53.3	26	75	Left parietal lobe and angular gyrus
GE	M	77	12	R (L)	3.8	3.0	71.0	89.7	75.0	40	60	Left parietal lobe
GA	M	50	18	R (R)	2.5	3.5	83.7	98.3	58.3	57	69	Left temporo-parietal and posterior Wernicke's area
RI	M	68	12	R (R)	3.4	2.0	84.2	84.3	56.7	49	60	Left temporal and fronto-parietal areas
CB	M	66	12	R (R)	5.8	4.0	90.0	93.3	86.7	49	58	Left temporo-occipital area, SM and angular gyri
ML	M	57	12	R (R)	7.5	3.5	92.5	97.3	68.3	53	68	Left SM and angular gyri, subcortical structures
LC	M	63	12	R (L)	5.8	4.0	90.0	93.3	86.7	49	69	Left posterior parietal lobe

It is important to note here that the classification of the aphasic patients was based on the original diagnosis made between one and six months post onset. All of the patients in this study have been followed clinically for years since their strokes. The Wernicke's patients all presented early on with classic neologistic jargon, paragrammatic output, and poor auditory comprehension on clinical testing. They have recovered to the point where their auditory comprehension scores are now on a par with the conduction aphasics; their language output still tended to be low in content words and paragrammatic, though neologisms were rare. Unfortunately, it is logistically

^a Includes overt and (familial) handedness.

impossible to analyze the coherence of more severe Wernicke's aphasics with neologistic output.

In addition to the aphasic patients, the study included 20 normal control subjects (11 females, 9 males) with no history of neurological disease, head injury, or active alcoholism. The control group was roughly matched with the aphasic patients for age and education. Among the aphasic patients, the mean age was 66.31 years (range 50-77 years) and the mean level of education was 11.92 years (range 8-18 years); among the normal control subjects, the mean age was 67.75 years (range 61-81 years) and the average level of education was 12.95 years (range 12-18 years). All subjects, normal and aphasic, were native speakers of American English.

Experiment one: Narrative Production

The purpose of the narrative production subtest of the RCB was to gain insight into the ability of aphasic patients to produce coherent stories, which included all of the propositions necessary to make the story complete but avoided irrelevant embellishments. The intrusion of irrelevant propositions in narrative production, as shown by Christiansen (1995), may indicate an inaccurate or underspecified macrostructure. The narrative production subtest in the current study is a partial replication of Christiansen (1995) with the additional inclusion of Broca's aphasics, a more detailed analysis of the range of relevant propositions, and a more carefully controlled rating system.

Stimuli and Data Analyses

The stimuli used for the narrative production subtest consisted of four cartoon stories of five frames each (one practice cartoon and three test cartoons). While the cartoons each contained a clearly depicted story, they also contained many additional background elements which were appropriate to the situational frame but irrelevant with regard to the story content. The purpose of these background items was to test the subject's ability to distinguish which items in the cartoons were relevant to the story and which items were not. The cartoon frames were prearranged in the correct order and presented in a vertical array. Each subject was asked to look at all of the pictures in the given cartoon, then tell the story that was depicted. All narrative samples were tape-recorded and transcribed verbatim for later analysis.

To determine the relevance of the propositions produced by both the

aphasic and normal groups, each narrative was submitted to the following three stages of analysis. First, the raw transcriptions were edited by eliminating self-corrections and abandoned utterances. Phonemic and morphological errors were ignored for the rest of the propositional analysis. Second, the edited texts were parsed into a series of semantic propositions based on a version of Kintsch's (1974) "text base" format, which was adapted to analyzing aphasic speech (see Christiansen, 1995). In general, the semantic propositions consisted of: 1) verbs and their arguments; 2) predicates and their arguments; and 3) adjunct adverbials of time, manner, and place. Only the story-based propositions (i.e., those propositions describing or elaborating on the cartoon pictures) were used for the relevance scoring. Personal comments made to the examiner about the story or task as well as repetitions and digressions were not scored. Third, all of the story-related propositions were ranked for degree of relevance to the main ideas of the stories by two independent raters, as described below.

Relevance rankings for the story-related propositions were based on a 7-point Likert scale. For each of the three test cartoon stories, a composite text base was compiled from all of the propositions given by both the normal and the aphasic patients, so that the raters were blind as to whether the propositions were given by an aphasic subject or a normal subject. The propositions on the master list were grouped according to picture frame and primary topic. The two independent raters then gave each proposition an overall ranking as to its relevance, which was based on how closely the proposition was connected to the main characters and plot of the story. A score of "7" indicated the most essential propositions; a score of "1" indicated completely irrelevant or illogical propositions, including confabulations.

After the raters had ranked all of the propositions, a master scoresheet was created for each cartoon story in which the rankings from the two raters were averaged for each proposition. The propositions in the text bases for each individual subject was then scored according to the master scoresheet. The scores for the propositions in each individual's story were averaged; the average ranks for each of the three test narratives were then summed and divided by 21 (the total possible if all propositions were ranked "7") to obtain a relevance score based on 100.

Results and Discussion

In general, the propositions produced by both the normal control group and the aphasic patients were ranked relatively high in relevance, however,

some significant differences did emerge. The normal subjects produced an average of 14.1 propositions per narrative and averaged 89.7 (s.d.=4.28) in their relevance scores. Responses ranking low in relevance (1-4) averaged 9.6% of the propositions produced, ranging from 0-21% for each subject. Most of the aphasic subjects could produce as many propositions per story as the normal subjects, averaging 11.96. Only the Broca's aphasics demonstrated a reduced output, each averaging less than 6.0 propositions per story. But despite producing a similar number of propositions, the aphasics' stories were generally ranked lower in relevance than the normal control subjects ($p=.02$; two-tailed T-test). As a group, the relevance scores of the aphasic patients averaged 83.2, with 20.2% of their propositions ranked between 1 and 4.

Figure 1 represents the individual z-scores from the narrative production test. As can be seen, each one of the Wernicke's aphasics scored at least three standard deviations below normal. On the other hand, all of the Broca's and conduction aphasics scored within normal limits, with the exception of one Broca's patient (RO), with a score of +2.41, indicating that he could only produce the most essential propositions for the story and lacked more peripheral elaborations.

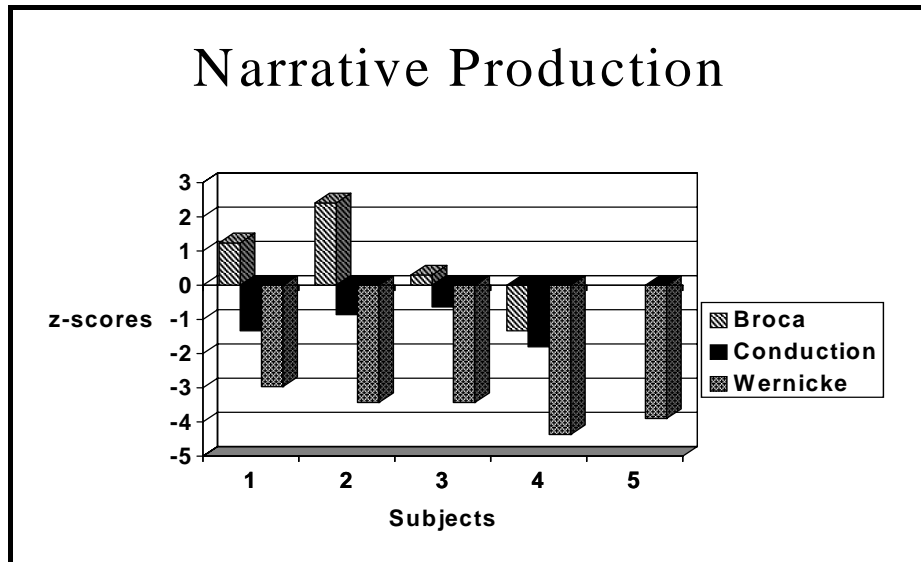


Figure 1. Individual z-scores for each of the aphasic patients on narrative production.

Experiment two: Sentence Pair Completion

In this subtest, the subject was required to form a probable scenario based upon a given lead-in sentence and produce a response sentence which would logically follow from the one given. As a follow-up to the narrative production test, the purpose in sentence pair completion was to determine if the aphasic patients could produce a relevant response when limited to just a single sentence. If the Wernicke's aphasics, who produced numerous irrelevant propositions in narrative production, also exhibited reduced relevance in their single sentence responses, this would indicate that the Wernicke's aphasics have a particular difficulty with judging even the local coherence between two sentences and that their difficulty in maintaining relevance may not be limited to connected discourse.

Stimuli and Data Analyses

The stimuli for this subtest consisted of 20 sentences and one practice sentence, each describing a real life problem. The examiner presented the sentence to the subject orally and asked the subject to respond by producing a logical follow-up sentence, as illustrated in item 1.

Examiner: James failed the math test, so ...

Subject: ... he studied harder for the next one.

All responses were tape-recorded and transcribed for later analysis. The transcribed responses were edited by removing false starts and phonological errors, as well as correcting grammatical errors. Grammatical errors were corrected so that the raters would be blind as to whether the responses were made by an aphasic patient or a normal subject.

The edited responses were then combined into a separate master list for each of the 20 stimulus sentences to be judged for their degree of relevance to the given situation. As in Experiment 1, the responses on each master list were put in random order, mixing both aphasic and normal responses, and each response was ranked by two independent raters using a 7-point Likert scale. A ranking of "7" indicated that the response was fully logical and required little processing effort on the part of the rater (e.g., James failed the math test, so *he studied harder next time*). A ranking of "4" could be viewed as logical, but either forced the rater to increase her processing effort to bridge the stimulus with the response by creating some possible intermediary scenario (e.g., to bridge „James failed the math test, so *he missed the party Saturday night*,“ the rater might imagine that James was punished for his

failure), or failed to provide adequate new information and simply rephrased information already implicit in the stimulus item (e.g., James failed the math test, so *he got zero*). A ranking of "1" indicated that the rater could not make logical sense of the response in the context of the stimulus sentence (e.g., James failed the math test, so *I'll be darned, he found it*).

To attain a relevance score for the sentence pair completion subtest, each subject's individual response was ranked according to the master scoresheet. If a subject produced more than one response for a particular stimulus item, each response was scored separately and averaged with the other responses for the stimulus item, so that each of the 20 items was given an average score between 1 and 7. The relevance scores for all of the items were then converted to a score based on 100, as in Experiment 1.

Results and Discussion

Performance of the normal control subjects on this task was almost equal to that on the narrative production task. Their average relevance score was 89.7 (s.d.=3.23). The number of low-relevance responses for the normal group averaged 5.25%, ranging from 0-20% for each subject. Among the aphasic patients, several had difficulty producing relevant follow-up responses in the sentence pair completion. Though the performance of the aphasic patients varied greatly as compared to the normal subjects, the aphasic group as a whole scored significantly lower in relevance ($p < .005$, two-tailed T-test), averaging only 78.7. While some patients produced no low-relevance responses, up to 60% of other patients' responses were ranked low in relevance.

When analyzing the performance of individual subjects, as seen in Figure 2, it first appears that the production of relevant verbal sequences is related to aphasic syndrome. However, results of a one-way ANOVA reveal no significant differences among the aphasic groups ($F(2,10) = 2.86$, $p = .10$). Judging from performance on the narrative production task, one would suspect the Wernicke's aphasics to have difficulty in sentence pair completion; in fact, the opposite appeared to be true. Only one Wernicke's patient (RI) had difficulty producing relevant sequences, scoring more than 7 s.d.'s below normal; the others were all well within normal limits. On the other hand, all of the Broca's and all but one of the conduction aphasics (PA) scored more than 2.5 standard deviations below normal.

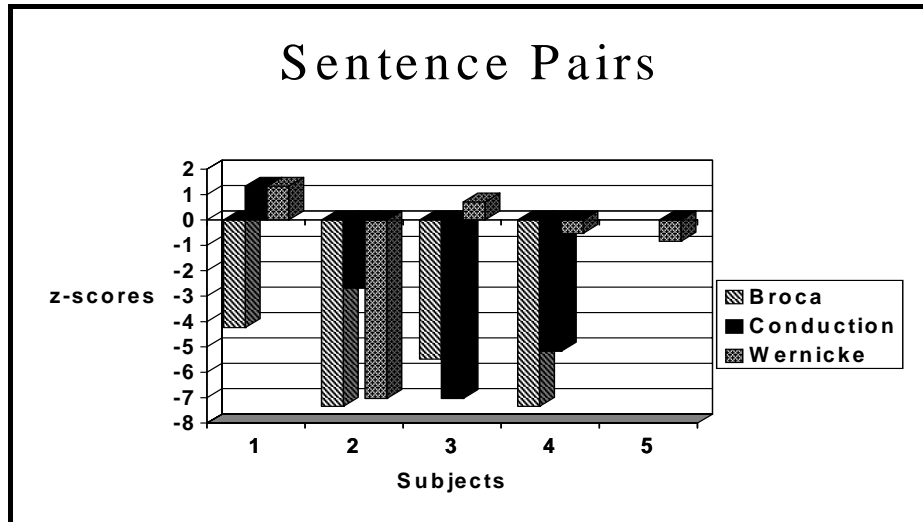


Figure 2. Individual z-scores for each of the aphasic patients on sentence pair completion.

Closer inspection, however, reveals that the apparent relation between relevance scores on sentence pair completion and aphasic syndromes may be misleading. Performance by the aphasic patients on the sentence pair completion task corresponded exactly with the patients' BDAE severity ratings. All of those patients who were given a severity rating of 3.0 or below scored at least two standard deviations below the normal control subjects in producing relevant sentence pairs; all patients with a severity rating of 3.5 or above scored within normal limits. The Wernicke's aphasics, in general, had higher severity ratings than the other two groups, probably due to the selection criteria. Only mild Wernicke's aphasics had recovered enough in auditory comprehension to score above the 70th percentile on the BDAE. Therefore, the subjects in the Wernicke's group performed well on sentence pair completion, with the exception of RI who had a severity rating of only 2.0. The Broca's and conduction aphasics, on the other hand, had severity ratings between 1.5 and 3.0, except for PA who had a 3.5. Future research would probably reveal that milder Broca's and conduction aphasics could also produce relevant responses on sentence pair completion, and that more severe Wernicke's would produce more irrelevant responses.

Experiment three: Word List Generation

The issue of processing load in maintaining topic relevance was most critically addressed in the word list generation subtest, as the subject only needed to produce single words without the added processing required to produce entire propositions and syntactic frames. In order to perform this task successfully, the subject had to evoke a mental image of a common setting or situation given by the examiner and verbally list items, events, or qualities usually associated with such a setting. The use of everyday settings such as schools or gas stations to elicit word lists, rather than lexical superordinates such as fruit or animals, was considered more closely related to narrative production, the only difference being that the subjects needed to access related lexical items arising from simple referring propositions rather than formulating complex propositions resulting in full sentences.

Stimuli and Data Analyses

In this test, the examiner presented the subject orally with a word or phrase describing a common script or frame (e.g., wedding, hospital, or gas station). The subject was then asked to say as many words as s/he could think of which would be associated with that particular setting. The responses could be nouns, verbs, adjectives or adverbs; in some cases subjects gave circumlocutory phrases (e.g., *machine for changing tires* for the gas station setting) and these responses were treated in the same way as single-word responses. Eight different topics were given in total. The subject was allowed ninety seconds to respond with as many words as possible, and all responses were tape-recorded for later transcription and analysis.

As in the first two experiments, a master list combining all of the responses given by both normal and aphasic subjects for each situation was constructed to analyze the relevance of each subject's responses. Phonemic errors were ignored and circumlocutions (e.g., *the thing you sit on and people push you* for "wheelchair") were given the same credit as the target word. All of the responses for each topic were simply listed alphabetically so that the raters were blind as to whether a particular response came from a normal or aphasic subject. Each response on the master list was then ranked, by two independent raters, on a 7-point Likert scale according to its centrality to the given setting. A "7" ranking meant that the item was a crucial component of the setting (e.g., *doctor* and *nurse* for "hospital"). A "4" ranking indicated that the word was commonly found within the setting, but

unnecessary or peripheral (e.g., *chair*, *cafeteria*, and *gift shop*, for "hospital"). A "1" ranking indicated an item which was either completely irrelevant or illogical within the given setting (e.g., *tires* for "hospital"). The rankings of the two raters were then averaged to create a master scoresheet. Individual scores were then calculated from the master list and converted to a 100-point scale.

Results and Discussion

The normal control subjects produced an average of 22.14 words (range 10.75-30.625) for each of the eight topics given, and their relevance scores averaged 87.6 (s.d.=2.80). The overall percentage of words produced which ranked low in relevance (1-4) was 7.5%, ranging from 3%-13%. The aphasic patients produced significantly fewer responses than the normals ($p < .001$, two-tailed T-test), averaging only 11.94 words per topic, but were comparable to the normal controls in relevance. Their overall relevance scores averaged 86.8, and though they produced a slightly higher proportion of low-relevance words, averaging 12.9% (range 4%-26%), this difference was not significant.

As can be seen in Figure 3, the relevance scores of the aphasic patients on the word list generation task were generally within normal limits, and individual difficulties were not related to any particular type of aphasia. One Broca's (MA), one conduction (CA), and one Wernicke's aphasic (RI) had z-scores less than -2.00, indicating that they produced significantly more words which were peripheral to a given topic. In contrast, one Broca's aphasic (MP) had a z-score of +2.32, indicating that she produced only the very central core words for a given topic. Deviant performance on this task does not appear to be related to overall severity of aphasia, as seen in Experiment 2, nor does it relate directly to naming difficulties as seen on the Boston Naming Test.

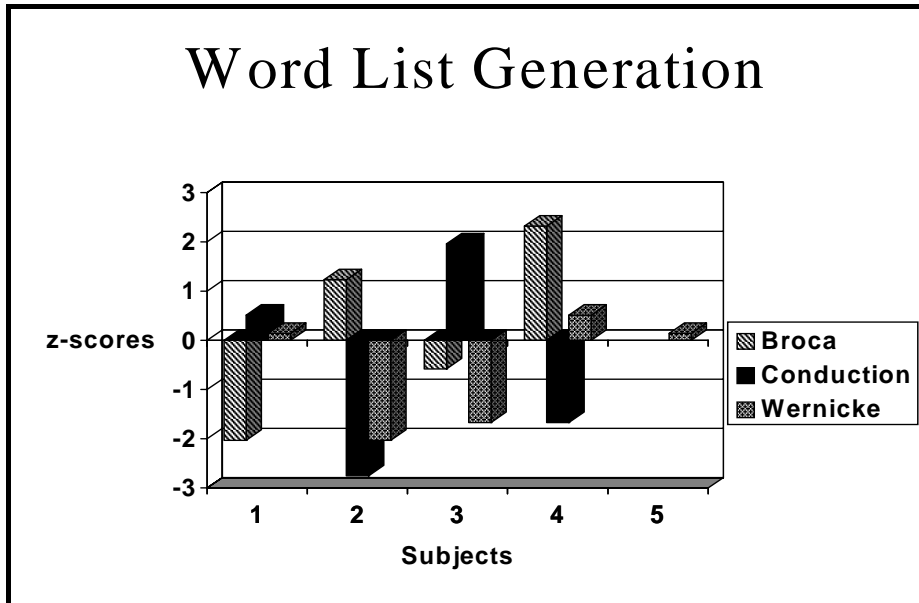


Figure 3. Individual z-scores for each of the aphasic patients on word list generation.

At first glance, the relevance scores of the aphasic patients on the word list generation task seem to indicate that their ability to construct situational frames was intact, as few subjects scored outside of the normal range. But although they produced words within a normal range of relevance, the aphasic patients produced only about half as many words per topic as the normal control subjects. This was a consistent pattern even among the very fluent aphasic patients. The combination of normal relevance with reduced lexical production would indicate that the aphasic subjects in general had not lost the more distally related concepts within a particular frame. Had these more peripheral concepts been lost, the aphasic patients would have produced fewer responses but would have demonstrated an increase in relevance ratings, as they would have only produced the most central features of the situational frame. Most of the aphasic patients in this study simply demonstrated an impoverished access to the necessary lexical items, regardless of their centrality to the situational frame.

Cross-task Comparison

Table 2 summarizes the relevance scores and z-scores for each of the aphasic subjects on all three tasks. To determine more precisely whether the Wernicke's aphasics have a selective disturbance in producing coherent narratives, the data from all three experiments were analyzed using a 3x3 ANOVA, comparing the three groups of aphasic patients across the three production tasks. Results from the two-way ANOVA revealed no main effect of group ($F(2,10) = 0.26, p=.78$), but did reveal a main effect of task ($F(2,10) = 5.35, p=.01$), as the aphasic patients scored lower on sentence pair completion as discussed in experiment two. The results also indicated a significant group x task interaction ($F(4,20) = 6.87, p=.001$), in which the Wernicke's aphasics scored significantly lower in relevance on the narrative production task than either the Broca's or conduction aphasics ($p<.01$ respectively on planned comparisons). To summarize, only the Wernicke's aphasics appear to have a limitation in producing coherent stories, but this limitation may prove to be particularly related to discourse and not a general inability to judge the relevance of information to a given topic.

Table 2
Aphasics' Individual Relevance Scores and Z-Scores

Subject	Narratives	Sentence Pairs	Word Lists
(Broca's)			
MA	95 (+1.24)	76 (-4.24)	82 (-2.03)
RO	100 (+2.41)	66 (-7.34)	91 (+1.23)
BL	91 (+0.30)	72 (-5.48)	86 (-0.58)
MP	84 (-1.33)	66(-7.34)	94 (+2.32)
(Conduction)			
PA	84 (-1.33)	94 (+1.33)	89 (+0.51)
CA	86 (-0.63)	81 (-2.69)	80 (-2.75)
GO	87 (-0.63)	67 (-7.03)	93 (+1.96)
GE	82 (-1.80)	73 (-5.17)	83 (-1.67)
(Wernicke's)			
GA	77 (-2.97)	94 (+1.33)	88 (+0.14)
RI	75 (-3.43)	67 (-7.03)	82 (-2.03)
CB	75 (-3.43)	92 (+0.71)	83 (-1.67)
ML	71 (-4.37)	88 (-0.53)	89 (+0.51)
LC	73 (-3.90)	87 (-0.84)	88 (+0.14)

Discussion

It should be stressed that the original purpose of this study was exploratory in nature, and the specific questions posed were: 1) whether aphasic patients could produce concise and relevant narratives, 2) whether difficulties in producing relevant discourse were associated with particular aphasic syndromes, and 3) whether the intrusion of irrelevant propositions in discourse might result from a general processing impairment also affecting the production of sentences and single words. To produce a coherent narrative from the cartoon picture stimuli, the subject first had to determine which parts of the cartoon pictures were crucial to the action series depicted and which parts were irrelevant embellishments of the setting. The subject would then base his/her narrative on the crucial actions and consequences in the cartoon. Difficulty in maintaining coherence in narrative production appears to be explicitly connected with Wernicke's aphasia, as every one of the Wernicke's aphasics produced an abnormal number of low-relevance propositions on the narrative production task, while none of the other aphasic patients scored outside of the normal range. These findings are consistent with Christiansen's (1995) previous finding, indicating that the strategy of determining the relevance of story elements may not be available to the Wernicke's aphasics. The results of the narrative production task further revealed that the Broca's aphasics performed as well as the conduction aphasics in terms of maintaining relevance, even though they produced fewer propositions.

Discourse and the Issue of Processing Load

If processing load has a direct effect on formulating relevant discourse at the micropropositional level, we would expect to find the following performance patterns among the aphasic patients: 1) difficulty only on the narrative production task, 2) difficulty on the sentence and discourse tasks but not on word list generation, or 3) difficulty on all three tasks. The difficulty maintaining coherence in narrative discourse as exhibited by the Wernicke's aphasics appears to be limited to discourse, however, the ability to maintain coherence may interact with the severity of the Wernicke's aphasia. RI, the most severe Wernicke's aphasic in the study, was the only subject who produced an abnormal number of irrelevant responses on all three production tasks. Unlike the Broca's and conduction aphasics, RI's irrelevant responses did not seem to be the result of

compensatory strategies, as some of his responses, especially on sentence pair completion and word list generation, showed little connection to the given topic, rather than simply lacking new information. In fact, it is possible that RI has lost part of his ability to monitor the relevance of his own utterances to the discourse context. Further research would have to determine whether RI's difficulty with producing relevant responses can be generally found in patients with a moderately severe Wernicke's aphasia or is an idiosyncratic disturbance.

An interesting and unexpected finding was that severity of aphasia affects the production of local coherence, as seen on the sentence pair completion task, but that this difficulty at the sentence level does not carry over to narrative production. The apparent relation between performance by the aphasic patients on the BDAE and their performance on the sentence pair completion test may seem logical at first, but one must stop and consider why a patient's ability to formulate language grammatically should affect his/her ability to produce a *relevant* response. It is likely that patients with more severe aphasia produced irrelevant responses not because of their overt language errors, but because of the alternative strategies they used to communicate. It is possible that aphasic patients were able to generate a relevant response at the prelinguistic level, but could not formulate it linguistically, so they attempted to generate a related response for which they could formulate a surface structure. However, the alternative responses were often less precise and circumlocutory, causing them to be rated lower in relevance. Less relevant responses often tended to be either very general, approaching automatic speech as illustrated in item 2, or lacking in new information, as illustrated in item 3.

Examiner: The situation in Bosnia is still pretty tense, so...

Patient: It's a mess. (rated 4)

Examiner: Too many children get hurt when their parents divorce, so...

Patient: They all get hurt. (rated 4)

Most of the aphasics' responses demonstrated some logical relation to the given sentence. Few patients gave responses which did not seem to fit the appropriate script being referred to by the examiner. Only RI (Wernicke's) and GE (conduction) produced illogical responses such as *The hell with the dog! Who cares about a dog?* in response to the stimulus sentence „My friend got her car stolen in Boston yesterday“, or *Take the elevator* in response to „Parking in the city is simply impossible“.

The obvious question that arises from the performance of the aphasic patients on sentence pair completion is: why should it be that severity of

aphasia affects the production of local coherence between single sentences but not global coherence in narrative discourse? One possibility is that for communication purposes, more severely aphasic patients benefit from the freedom of a discourse context, since they are then allowed to produce as many utterances as they wish to get their point across. When restricted to a single response utterance, more severely impaired aphasic patients do not seem to be able to embed enough information into one utterance, and therefore their responses were judged as less relevant than the more mildly impaired aphasic patients who had a greater capacity to embed information. Such a notion would agree with Goodglass et al.'s (1994) previous finding that while normal age-matched subjects produced an average ratio of approximately two propositions per utterance in free speech, both Broca's and conduction aphasics could only produce a roughly one-to-one ratio. While the Broca's patients in Goodglass et al.'s study produced the same number of utterances as the normal subjects, thereby producing only half as many propositions, the conduction aphasics made up for their lack of embedding by producing twice as many utterances as the normal group, therefore producing approximately the same number of propositions. This same production behavior was found in the narrative production task of the current study. The Broca's patients may also have difficulty producing coherent narratives, however, their lack of coherence results from a lack of essential information in their narratives, and not from a lack of relevance relation between their story propositions and the main points of the intended story, as discussed below.

Discourse Production in Broca's Aphasia

From the beginning of the 20th century, aphasiologists supporting the economy of effort hypothesis espoused that agrammatic Broca's aphasics restricted their language output to only the most essential elements of their intended message. Pick (1913) expressed the notion of „Notsprache“ (emergency speech) as a strategy in the discourse of agrammatic Broca's patients in which the patients try to convey the most core information in the shortest possible time. Although the idea of Notsprache was used to explain the telegraphic nature of agrammatic speech, it also related to the overall sparse language production of many Broca's aphasics. In the current study, the use of Notsprache should be evident in the relevance scores of the Broca's aphasics on narrative production and word list generation, where production was not limited to a single response. We would expect to find

unusually high relevance ratings, more than two standard deviations above normal. Such a finding would indeed indicate that the Broca's aphasics were attempting to convey only the words or propositions most central to a given topic.

Evidence for the economy of effort strategy by the Broca's aphasics in the current study was far from overwhelming. Although two patients performed above the normal range on one task each, RO scoring +2.41 on narrative production and MP scoring +2.32 on word list generation, all other relevance scores by the Broca's aphasics fell within the normal range, including both core words or propositions as well as concepts more peripheral to the topic. In narrative production, all of the conduction aphasics produced more propositions per story than any of the Broca's aphasics, yet they were not less relevant. The Broca's aphasics tended to omit key propositions, while including peripheral propositions. Even though the Broca's patients produced fewer words per topic in word list generation than any of the fluent aphasics, except for GO (conduction), they included such peripheral words as *turnips* (BL) or *cold* (MP) under the Restaurant topic, *stones* (BL) or *parking the car* (RO) under the City Park topic, and produced *donkey* but not *cow* (MA) under the Farm topic. The reduced output of the Broca's aphasics in conjunction with their maintenance of a normal range of essential and optional information in their output would seem to rule out a production strategy based on economy of effort. It remains an open question as to why the Broca's aphasics leave out essential information in their discourse even at the propositional level, and this global reduction in output should be considered in future theories of agrammatic production.

Conclusion

To summarize the overall findings of the current study, Wernicke's aphasics, and only Wernicke's aphasics, consistently have difficulty producing coherent story narratives without the intrusion of irrelevant propositions. This difficulty appears only to affect production at the discourse level, however, there may be an interaction with the severity of the Wernicke's aphasia. Unfortunately, the interaction between Wernicke's aphasia and the severity of the aphasia will be extremely difficult to investigate, given the frequent neologistic output and attentional limitations of more severely impaired Wernicke's patients. For other patients, severity of aphasia only affects the relevance of their responses when their output is limited by task demands, such as restricting them to only one sentence. An additional finding is that even though the language output of the Broca's

aphasics is severely reduced as compared to normal control subjects, they nonetheless produce a similar proportion of essential and peripheral concepts when given a specific topic. It must be stressed that the current study was purely exploratory in nature and that the findings should form the basis for future research questions rather than providing definitive answers about aphasic patients' ability to produce relevant language output.

References

- Bihrlle, A.M., Brownell, H.H., Powelson, J.A., & Gardner, H. (1986). Comprehension of humorous and nonhumorous materials by left and right brain-damaged patients. *Brain and Cognition*, 5, 399-411.
- Brownell, H.H. & Gardner, H. (1988). Neuropsychological insights into humour. In J. Durant and J. Miller (Eds.), *Laughing matters: A serious look at humour* (pp. 17-34). New York: John Wiley & Sons, Inc.
- Brownell, H.H., Michel, D., Powelson, J., & Gardner, H. (1983). Surprise but no coherence: Sensitivity to verbal humor in right-hemisphere patients. *Brain and Language*, 18, 20-27.
- Brownell, H.H., Potter, H.H., Bihrlle, A.M., & Gardner, H. (1986). Inference deficits in right brain-damaged patients. *Brain and Language*, 27, 310-321.
- Brownell, H.H., Potter, H.H., Michelow, D., & Gardner, H. (1984). Sensitivity to lexical denotation and connotation in brain-damaged subjects: A double dissociation? *Brain and Language*, 22, 253-265.
- Brownell, H.H., Simpson, T.L., Bihrlle, A.M., Potter, H.H., & Gardner, H. (1990). Appreciation of metaphoric alternative word meanings by left and right brain-damaged patients. *Neuropsychologia*, 28 (4), 375-383.
- Christiansen, J.A. (1993, January). *Coherence disturbance in Wernicke's aphasia*. Paper presented at the Annual Meeting of the Linguistic Society of America, Los Angeles, California.
- Christiansen, J.A. (1994). *Coherence disturbances in aphasia: The assessment of relevance*. Unpublished doctoral dissertation, Boston University, Boston.
- Christiansen, J.A. (1995). Coherence violations and propositional usage in the narratives of fluent aphasics. *Brain and Language*, 51, 291-317.
- Clark, H.H. (1975). Bridging. Reprinted in P.N. Johnson-Laird and P.C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 411-420).

- Cambridge, England: Cambridge University Press, 1977.
- Clark, H.H. & Haviland, S.E. (1977). Comprehension and the given-new contract. In R.O. Freedle (Ed.), *Discourse production and comprehension* (pp. 1-40). Norwood, New Jersey: Ablex Publishing Corp.
- De Renzi, E. & Faglioni, P. (1978). Normative data and scoring power of a shortened form of the Token Test. *Cortex*, 15, 41-49.
- Foldi, N.S. (1987). Appreciation of pragmatic interpretations of indirect commands: Comparison of right and left hemisphere brain-damaged patients. *Brain and Language*, 31, 88-108.
- 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.
- Glosser, G. & Deser, T. (1990). Patterns of discourse production among neurological patients with fluent language disorders. *Brain and Language*, 40, 67-88.
- Goodglass, H., Christiansen, J.A., & Gallagher, R. (1994). Syntactic constructions used by agrammatic speakers: Comparison with conduction aphasics and normals. *Neuropsychology*, 8(4), 598-613.
- Goodglass, H. & Kaplan, E. (1983). *Boston Diagnostic Aphasia Examination* (rev. ed.). Hillsdale, New Jersey: Lea & Febiger.
- Goodglass, H., Kaplan, E., & Weintraub, S. (1983). *Boston Naming Test*. Hillsdale, New Jersey: Lea & Febiger.
- Gordon, P.C. (1993). Computational and psychological models of discourse. In H. Brownell and Y. Joanette (Eds.), *Narrative discourse in neurologically impaired and normal aging adults* (pp. 23-46). San Diego: Singular Publishing Group, Inc.
- Grosz, B.J. & Sidner, C.L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3), 175-204.
- Haviland, S.E. & Clark, H.H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.
- Hirst, W., Ledoux, J., & Stein, S. (1984). Constraints on the processing of indirect speech acts: Evidence from aphasiology. *Brain and Language*, 23, 26-33.
- Huber, W. (1990). Text comprehension and production in aphasia: Analysis in terms of micro-and macroprocessing. In Y. Joanette and H. H. Brownell (Eds.), *Discourse ability and brain damage* (pp. 154-179). New York: Springer Verlag.

- Joanette, Y. & Goulet, P. (1990). Narrative discourse in right-brain-damaged right-handers. In Y. Joanette and H. H. Brownell (Eds.), *Discourse ability and brain damage* (pp. 131-153). New York: Springer Verlag.
- Joanette, Y., Goulet, P., Ska, B., & Nespoulous, J.-L. (1986). Informative content of narrative discourse in right-brain-damaged right-handers. *Brain and Language*, 29, 81-105.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Kintsch, W. (1991). The role of knowledge in discourse comprehension: A construction-integration model. In G. Denhiere and J.-P. Rossi (Eds.), *Text and text processing* (pp. 107-153). Amsterdam: Elsevier Science Publishers B.V.
- Kintsch, W., & Van Dijk, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85 (5), 363-394.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, Massachusetts: MIT Press.
- Molloy, R., Brownell, H., & Gardner, H. (1990). Discourse comprehension by right-hemisphere stroke patients: Deficits of prediction and revision. In Y. Joanette and H. H. Brownell (Eds.), *Discourse ability and brain damage* (pp. 113-130). New York: Springer-Verlag.
- Myers, P.S. (1993). Narrative expressive deficits associated with right-hemisphere damage. In H. Brownell and Y. Joanette (Eds.), *Narrative discourse in neurologically impaired and normal aging adults* (pp. 279-296). San Diego: Singular Publishing Group, Inc.
- Pick, A. (1913). *Die agrammatischen Sprachstörungen*. Berlin: Karger.
- Sandson, J., Albert, M., And Alexander, M.P. (1986). Confabulation in aphasia. *Cortex*, 22, 621-626.
- Sperber, D. & Wilson, D. (1986). *Relevance: Communication and cognition*. Cambridge, Massachusetts: Harvard University Press.
- Ulatowska, H.K., Allard, L., & Chapman, S.B. (1990). Narrative and procedural discourse in aphasia. In Y. Joanette and H.H. Brownell (Eds.), *Discourse ability and brain damage* (pp. 180-198). New York: Springer-Verlag.
- Ulatowska, H.K., Freedman-Stern, R., Doyel, A.W., Macaluso-Haynes, S., & North, A.J. (1983). Production of narrative discourse in aphasia. *Brain and Language*, 19, 317-334.
- Ulatowska, H.K., North, A.J., & Macaluso-Haynes, S. (1981). Production of narrative and procedural discourse in aphasia. *Brain and Language*, 13, 345-371.

- Van Dijk, T.A. (1977). *Text and context. Explorations in the semantics and pragmatics of discourse*. London: Longman.
- Van Dijk, T.A. (1980). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Van Dijk, T.A. & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.
- Van Lancker, D.R. & Kempler, D. (1987). Comprehension of familiar phrases by left- but not by right-hemisphere damaged patients. *Brain and Language*, 32, 265-277.
- Weylman, S.T., Brownell, H.H., & Gardner, H. (1988). "It's what you mean, not what you say": Pragmatic language use in brain-damaged patients. In F. Plum (Ed.), *Language, communication, and the brain* (pp. 229-243). New York: Raven Press.
- Weylman, S.T., Brownell, H.H., Roman, M., & Gardner, H. (1989). Appreciation of indirect requests by left- and right-brain-damaged patients: Effects of verbal context and conventionality of wording. *Brain and Language*, 36, 580-591.
- Winner, E. & Gardner, H. (1977). The comprehension of metaphor in brain-damaged patients. *Brain*, 100, 717-729.

Received April 06, 2000
Revision received May 01, 2000
Accepted May 23, 2000