

## ABSTRACT

Title of Dissertation / Thesis: A PILOT STUDY TO DEVELOP DISCOURSE  
CODES SPECIFIC TO PREFRONTAL  
DYSFUNCTION

Inbal Eshel, Master of Arts, 2004

Dissertation / Thesis Directed By: Nan Ratner, Ed.D., Chair, Department of  
Hearing and Speech Sciences

This pilot study developed a set of codes designed to capture the “nonaphasic” but characteristic discourse deficits that may be present following prefrontal cortex damage (PFCD). The codes were utilized based on narrative sample elicitation to investigate between-group differences in two study populations: patients with left, right, or bi-frontal PFCD and age and education-matched healthy comparison group participants. Narrative samples were coded on indices of content units, thematic units, story grammar features, and discourse errors, and analyzed using CLAN. Results of this study support the original deficit hypotheses. The coding schema demonstrated fair to good inter-rater reliability, stronger performances by the healthy comparison group across all four levels of analysis, and poorer performance overall on the retell phase than the tell phase. Qualitative analysis revealed relatively few discourse errors associated with the healthy comparison group, while various classic discourse errors were associated with the PFCD group.

A PILOT STUDY TO DEVELOP DISCOURSE CODES SPECIFIC TO  
PREFRONTAL DYSFUNCTION

By

Inbal Eshel

Thesis submitted to the Faculty of the Graduate School of the  
University of Maryland, College Park, in partial fulfillment  
of the requirements for the degree of  
Master of Arts  
2004

Advisory Committee:  
Professor Nan Ratner, Ed.D, Chair  
Carol Frattali, Ph.D †  
Rochelle Newman, Ph.D  
Froma Roth, Ph.D

© Copyright by  
Inbal Eshel  
2004

## Dedication

To Carol.

I can still feel your support.

I can still feel you pushing me to reach my fullest potential,  
encouraging my independence and growth.

I can still feel your insight, your brilliance, your wisdom, your  
strength.

I can still feel your pride, in our shared accomplishment and in my  
personal achievements.

Your passions became mine; you inspired me more than I could have  
imagined. You have left your mark on my thoughts and aspirations.

To my dear friend, mentor, advisor –

This is for you.

## **Acknowledgments**

Dan, Ronit, Itzik, and Neir – you are my support, you are my family. I have no words....

To Nan – for your enduring care, nurturing, support, and belief in me.

To Dr. Roth and Dr. Newman – for all of your advice and encouragement.

A big thank you to Yun and Mary, for all your help, in so many different ways.

A special thanks to Angelique LaPeyre and Laura Gutowski for volunteering to act as reliability judges.

# Table of Contents

Dedication .....	ii
Acknowledgments.....	iii
Table of Contents.....	iv
List of Tables .....	vi
List of Figures .....	vii
INTRODUCTION .....	1
Literature Review.....	1
Discourse, defined.....	1
Discourse Impairments in Acquired Language Disorders .....	3
Effects of Prefrontal Cortex Damage on Discourse.....	5
A Model of Prefrontal Function Relevant to Discourse .....	8
Narrative Discourse Production Tasks.....	10
Discourse Impairments in Developmental Language Disorders .....	12
Methods of Discourse Production Analysis.....	14
Rationale for the current study.....	17
Method Overview .....	18
PHASE I.....	19
Purpose.....	19
Piloting of the Narrative Discourse Task.....	20
Specific Task Details .....	21
Results.....	23
Task Revisions, Based on Pilot Results.....	27
PHASE II .....	28
Participants.....	28
Experimental Stimuli .....	33
Procedure .....	34
Specific Task Details .....	34
Data Analysis and Reduction.....	36
Transcription of Sample/Coding.....	36
Selection of Final Codes for Data Analysis.....	39
Data Analysis.....	40
Results.....	40
Do the Discourse Codes Demonstrate Sufficient Inter-rater Reliability?.....	41
Post-hoc Analysis Procedure: Lexical Efficiency.....	43
Descriptive Statistics.....	46
Between-group Differences: Story Processing Time, Comprehension Question Accuracy, Content Units, Thematic Units, Story Grammar Features, and Discourse Errors.....	47
Between Group Differences – Tell vs. Retell .....	50
Qualitative Analysis –Between-group Differences: Story Grammar .....	52
Qualitative Analysis – Summary of Data Trends .....	54
DISCUSSION .....	55

Overview.....	55
Do the Codes Demonstrate Sufficient Inter-rater Reliability?.....	57
Reliability – Suggestions Relative to the Training Process.....	58
Reliability – Suggestions Relative to the Task.....	59
Between-group Differences: Content Units, Thematic Units, Story Grammar Features, and Discourse Feature Errors.....	59
Between-group Differences – Tell vs. Retell.....	61
Qualitative Analysis – Between-group Differences.....	64
Discourse Models as Applied to Results.....	67
Study Limitations.....	68
Challenges in Using Narrative Analysis as a Clinical Tool.....	69
Suggested Task Adjustments.....	70
Study Results in the Context of Current Literature.....	72
Implications for Future Research.....	73
CONCLUSION.....	74
APPENDIX A – Discourse Analysis Systems.....	75
Automated Systems.....	75
CLAN.....	75
SALT.....	77
Manual Systems.....	77
Cohesion Analysis.....	79
Intonation unit analysis of conversational discourse.....	80
Linguistic Communication Measure.....	81
Quantitative Analysis of Aphasic Sentence Production.....	82
Systemic Functional Linguistics.....	84
Topic Analysis.....	85
Appendix References.....	86
APPENDIX B – Discourse Error Analysis.....	88
Appendix References.....	90
APPENDIX C – Pilot Questionnaire.....	92
APPENDIX D – Compilation of Pilot Questionnaire Comments.....	93
APPENDIX E – Instructions/Pictorial Stimuli/Comprehension Questions.....	94
APPENDIX F – Reliability Training Manual.....	101
WORKS CITED.....	111

## List of Tables

Table 1. Number and percentage of participants according to age group by gender, education level, and race/ethnicity.....	20
Table 2. Number and percentage of participants who referenced frame content in story tell according to age by frame number and description. ....	23
Table 3. Number and percentage of participants who referenced thematic content unit according to age by thematic unit number and description. ....	24
Table 4. Percentage of correct responses according to age by question number.....	25
Table 5. Number and percentage of participants according to age by responses to questionnaire. ....	26
Table 6. Demographic Data, Reaction Times, and Comprehension Question Accuracy, by Group.....	29
Table 7. Mean Lesion Size, by Group .....	30
Table 8. Test Score Data, by Group.....	33
Table 9. Kappa Statistic for Inter-Rater Reliability .....	42
Table 10. Intraclass Correlation Coefficient – Total Discourse Errors .....	43
Table 11. Descriptive Statistics: Content Units (CU), Thematic Units (TU), and Story Grammar (SG) .....	47
Table 12. Discourse Error Analysis .....	50
Table 13. Story Grammar Counts by Category .....	53
Table 14. Discourse errors not present in either group’s sample.....	55
Table 15. Excerpts from participant transcripts.....	64



## List of Figures

Figure A. Story Grammar Change by Group.....	51
Figure B. Content Unit Change by Group .....	52
Figure C. Thematic Unit Change by Group.....	52

## **INTRODUCTION**

A complex system of cognitive and linguistic processes underlies the everyday use of language. Language can be viewed and analyzed on many levels, one of which is “language in use” (Frattali & Grafman, in press), or discourse. Compared to production of sounds, words, or sentences in isolation, discourse production as an integrative and context-driven construct is thought to be representative of the complex communication needed for daily life activities. Therefore, cognitive and linguistic analysis at the level of discourse should be more sensitive to characterizing the types of communication deficits that various clinical populations may exhibit in the context of daily living.

## **Literature Review**

### **Discourse, defined**

Discourse can be defined broadly as language use “in the large,” (Clark, 1994, p. 985), or as extended activities that are carried out via language (Clark, 1994).

Discourse can be explored either at the level of comprehension or production (Brown & Yule, 1983; Caplan, 1999; Clark, 1994; Frattali & Grafman, in press). In terms of receptive skills, discourse processing refers to the ability to establish relationships within and between sentences, using context as the foundation for comprehension to form a coherent representation (Brown & Yule, 1983; Frattali & Grafman, in press).

In terms of expressive ability, discourse production can be transactional or interactional. Transactional discourse refers to the expression of content, while interactional discourse refers to the expression of personal attitudes and social

relationships (Brown & Yule, 1983; Frattali & Grafman, in press). Discourse can also be examined via a text view (e.g., discourse as a product) or as a joint activity (e.g., discourse as a process). Because of its inherently dyadic nature, Clark (1994) suggests that it is more meaningful to view discourse as a joint activity, which applies to interactional conversation as well as to stories told to others by single narrators. In the latter scenario, Clark (1994) notes that the listener is involved, albeit in a more passive role, in creating a mental representation of the narrative world fashioned by the narrator.

When investigating discourse, researchers have also often distinguished microlinguistic from macrolinguistic abilities (Glosser & Deser, 1990; Ulatowska, North, & Macaluso-Haynes, 1981; Ulatowska, Freedman-Stern, Doyle, & Macaluso-Haynes, 1983). Microlinguistic abilities refer to the processing of phonological, lexical-semantic, and syntactic aspects of single words and sentences. Measures of syntactic complexity and production at the single word level are often used to tap microlinguistic abilities (Glosser & Deser, 1990). Macrolinguistic abilities refer to the maintenance of conceptual, semantic, and pragmatic organization at the suprasentential level (Cannizzaro & Coelho, 2002; Glosser & Deser, 1990). Macrostructure relies on the interaction of both linguistic and non-linguistic knowledge, especially the non-linguistic systems of executive control and working memory (Cannizzaro & Coelho, 2002).

Coherence and cohesion are often used as measures of macrolinguistic abilities (Halliday & Hasan, 1976). Coherence refers to the ability to maintain thematic unity, and can be quantified as “global” (overall organization of goal, plan,

or theme; Glosser & Deser, 1990) or “local” (links between individual propositions or sentences which help maintain conceptual meaning; Glosser & Deser, 1990).

Cohesion refers to specific “relations of meaning between elements within discourse” (Glosser & Deser, 1990, p. 70). For example, anaphoric cohesion consists of linking a pronoun back to its reference (e.g., Bob’s home. He just walked in the door).

### **Discourse Impairments in Acquired Language Disorders**

Much research on discourse explores whether microlinguistic and macrolinguistic abilities can be dissociated neurologically and psychologically. These abilities have been investigated in various clinical populations, including traumatic brain injury (TBI) and right-hemisphere brain-damaged patients (RHBD) (Cannizzaro & Coelho, 2002; Coelho, 2002; Davis, O’Neil-Pirozzi, & Coon, 1997; Glosser & Deser, 1990; Joannette, Goulet, Ska, & Nespoulous, 1986; Mentis & Prutting, 1987; Togher & Hand, 1999; Tucker & Hanlon, 1998).

Research has explored the macrolinguistic and microlinguistic narrative discourse production abilities of TBI survivors. The literature suggests that, compared to non-brain-injured (NBI) controls, TBI survivors evidence impairment in macrolinguistic abilities, producing discourse that contains less output (Coelho, 2002) and contains deficits in coherence and cohesion (Cannizzaro & Coelho, 2002; Glosser & Deser, 1990; Mentis & Prutting, 1987). Their discourse also contains fewer implied meanings and is more concrete (Tucker & Hanlon, 1998), with more pragmatic errors (Snow, Douglas, & Ponsford, 1999) than NBI controls. In terms of microlinguistic abilities, their discourse also contains a greater number of syntactic

and lexical errors (Glosser & Deser, 1990) than NBI controls. Overall, TBI survivors demonstrate both microlinguistic and macrolinguistic deficits in discourse production.

The literature suggests that RHBD patients present primarily with macrolinguistic deficits. In terms of expressive language, McDonald (2000) describes RHBD patients as tangential, inefficient, and verbose, as well as impaired in inferencing skills. In addition, pragmatic impairments are often noted in RHBD patients, including inappropriate speech act use and interpretation, lack of sensitivity to situation and listener needs, and literal interpretation of figurative and implied meanings (Tompkins, 1995).

In terms of discourse comprehension, lesion studies have produced considerable evidence suggesting that adults with RHBD have difficulty drawing inferences. It has been suggested that the right hemisphere specifically contributes to discourse comprehension more than to single word comprehension (Beeman, 1993). The right hemisphere may also play a critical role in revising interpretations and building organized mental structures to form a mental representation of discourse (Beeman, 1993). Discourse impairments may be due, in part, to ineffective suppression of contextually irrelevant or inappropriate meanings (Tompkins, Baumgaertner, Lehman, & Fossett, 1997). Therefore, impairment may also be related to difficulty combining information across sentences, despite preserved processing of individual sentences.

Wapner, Hamby, and Gardner (1981) found that some RHBD patients are poor at inferring motives and morals from story contexts. Some RHBD patients may also experience difficulty integrating the elements of a story into a coherent narrative.

This process may be disrupted by the interjection of personal references, rationalization of foreign elements, and confabulation (Moya, Benowitz, Levine, & Finklestein, 1986; Wapner et al., 1981). Overall, research suggests that RHBD patients experience deficits at the supra-sentential, or macro, level of discourse. Much research describes RHBD as a unitary phenomenon, with little description of topographic representation within the hemisphere as related to its role in discourse processing. To state it differently, the right hemisphere is often described as a whole, without specifying distinct regions that may contribute to various aspects of discourse. This suggests that it is possible that regardless of the precise region of damage, impaired right hemisphere function contributes to discourse-related difficulties.

### **Effects of Prefrontal Cortex Damage on Discourse**

Damage to either Broca's or Wernicke's areas is traditionally associated with drastic changes in language ability. In contrast, focal cortical damage and its interruption to subcortical pathways can also be associated with more subtle changes in language. One population of patients, namely those with prefrontal cortex damage (PFCD), often present with such subtle language deficits that nevertheless can have a profound effect on functional communication. The prefrontal cortex is that portion of the frontal lobe anterior to the motor strip, and can be subdivided into dorsolateral (Brodmann's areas 8, 9, 10, and 46), orbitofrontal (Brodmann's areas 10, 11-13, and 47), and medial frontal/cingulate [Brodmann's areas 6, 8-10, 12, and 23, 24, and 32 (anterior cingulate)] areas.

Traditionally, the PFC is often described as mediating the cognitive processes of short-term and working memory, preparatory set, and inhibitory control (Fuster, 1997), as well as action planning (Alexander, 2002) and attention (Ferstl, Guthke, & Cramon, 1999). Classically thought to be non-specific to language use, some researchers suggest that many of the subtle language deficits exhibited following PFCD may in fact be a consequence or symptom of primary cognitive deficits (e.g., Ferstl et al., 1999). Specifically, these cognitive deficits may include action planning (Alexander, 2002), memory, and attention (Ferstl et al., 1999). However, Frattali and Grafman (in press) note that findings from neuroimaging studies suggest that attributing the full range of language deficits post-PFCD only to cognitive dysfunctions may be misguided. In fact, the PFC may have a specific role in context-sensitive semantic processing and selection (Binder, Frost, Hammeke, Cox, Rao, & Prieto, 1997; Demb, Desmond, Wagner, Vaidya, Glover, & Gabrieli, 1995; Frattali & Grafman, in press; Kapur, Rose, Liddle, Zipursky, Brown, Stuss, Houle, Tulving, 1994; Poldrack, Wagner, Prull, Desmond, Glover, & Gabrieli, 1999). Showing a specific linguistic role for the PFC, Demb et al. (1995) and Kapur et al. (1994) both found greater activation in the left PFC in semantic tasks relative to non-semantic tasks. In addition, Poldrack et al. (1999) found evidence for functional specialization of semantic and phonological processing in the left inferior prefrontal cortex.

Despite evidence of impairments to language following damage to the PFC, the language and discourse deficits of patients with PFCD are often considered to be difficult to document on traditional clinical tests, while profound in their damaging effect on daily life routines. PFCD patients are sometimes termed “nonaphasic”

because the classic characteristics associated with aphasia (e.g., syntactic impairments, lexical errors, and deficits in auditory comprehension) are relatively uncommon in this population. As described by Frattali and Grafman (in press), communication with a person with PFCD leads to impressions of perceived missteps in conversation, while responses “may fall short of what is considered expectable, acceptable, or sufficient from pragmatic, propositional, or contextual points of view” (p. 2).

PFCD patients have also been described as having largely intact microlinguistic abilities of word- and sentence-level processes, with impaired suprasentential, “text level” function (Ferstl et al., 1999). Overall, there have been considerably fewer studies describing the effects of prefrontal cortex damage on discourse than damage to other cortical areas (Alexander, 2002). However, the literature suggests that characteristics associated with PFCD patients’ discourse production include failure to stay within a given topic, tangentiality, lack of cohesion, difficulties with temporal sequencing, and reduced or enhanced speech output (Ferstl et al., 1999). In addition, in one study, McDonald (1993) describes striking similarities between the language impairments seen after RHBD to those seen after PFCD. These common discourse impairments include: verbosity, disorganization, tangentiality, concreteness, and an inability to interpret or utilize conversational inference (McDonald, 1993). However, overall, little exploration of these deficits has occurred to date.

Patients with PFCD may experience various difficulties in the context of storytelling. These include difficulty recalling narrative components of a story,



processing inference, and appreciating the story's thematic aspects or gist (Zalla et al., 2002; Frattali & Grafman, in press). Other characteristics of discourse deficits that may be demonstrated on a story tell/retell task include confabulation, embellishment, topic stray, ambiguous statements, faulty anaphoric reference and links, and faulty temporal sequencing of events and cause/effect relations (Craig & Frattali, 2000; Frattali & Grafman, in press). Several other deficits resulting from PFC may include loss of moralistic meaning (Zalla et al., 2002), misinterpreting abstract or implicit information, and producing a story tell/retell that either contains intrusive detail or is lacking in detail.

### **A Model of Prefrontal Function Relevant to Discourse**

Grafman's framework of PFC function, the Structured Event Complex (SEC), views the PFC as representational, and not purely process-oriented. In a process-oriented approach, the PFC manipulates information that is stored elsewhere in the brain (Wood & Grafman, 2003). Other process models, such as models of working memory, include descriptions of PFC performance; however, Grafman (2002) claims that they do not adequately account for the underlying representation that is responsible for the performance. In contrast to process-oriented approaches, representational approaches specify the type of information that is stored in memories in the PFC (Wood & Grafman, 2003).

Grafman's (2002) representational framework is structured around the SEC itself, as a basic processing unit, which he defines as a "set of events, structured in a particular sequence, that as a complex composes a particular kind of activity that is usually goal oriented" (p. 298). Grafman (2002) claims that the PFC contains

multiple components of higher-level knowledge, and that these are distinctive memory domains. To strengthen the role of the involvement of the PFC in higher-level knowledge, Wood and Grafman (2003) note that the pyramidal cells in the PFC are structurally equipped to handle more excitatory input than other cortical pyramidal cells, which is one possible explanation for the PFC's ability to integrate input from many sources and to implement more abstract behaviors. In addition, Grafman (2002) suggests that different knowledge forms are stored topographically (in distinct regions) within the PFC. For example, subsequent to damage to the ventromedial PFC, impairment of social behaviors (e.g. social rules, attitudes, scripts, etc.) appear to be most evident (Wood & Grafman, 2003). In contrast, subsequent to dorsolateral PFC damage, reflective, mechanistic behavior appears to be impaired (Wood & Grafman, 2003).

Grafman (2002) provides additional evidence of an SEC-type representational network, including the representation of several different SEC components within the PFC. In two examples cited by Grafman (2002), both event sequencing (Sirigu, Zalla, Pilon, Grafman, Dubois, & Agid, 1995) and thematic knowledge (Zalla, Phipps, & Grafman, 2002) can be impaired with damage to the PFC, even when event knowledge is almost or completely preserved. Consistent with the SEC model, it can be inferred that the PFC mediates temporal sequencing of events, cohesion, coherence, and gist, and that PFCD should produce predictable errors in narrative discourse tasks.

## **Narrative Discourse Production Tasks**

Narrative production tasks provide an important tool in the analysis of discourse deficits. The nature of narrative production tasks bears critically on differences between groups of subjects. Narrative discourse production tasks have demonstrated subtle, often complex communication deficits exhibited by patients who have suffered a traumatic brain injury (Coelho, 2002; Tucker & Hanlon, 1998). These communication deficits include difficulty remaining within a given topic, difficulty recalling narrative components of a story, failure to structure, misinterpretation of abstract or implicit information, and difficulty in planning and sequencing.

Discourse can be elicited in many ways. Narrative tasks (monologues) are often associated with increased communicative responsibility, making them more complex than dialogues (conversation). For example, a successful narrative relies on the coordination of many aspects within the narrator, including spatial and temporal frames of reference, individual objects, states, events, and processes relating to the speaker's frame of reference, the experience of changes in the objects, states, events, and processes as the situation unfolds, etc. (Clark, 1994). Coordination is also necessary in conversation, but it is shared among multiple speakers; they share the communicative responsibility. Overall, a successful narrative requires the narrator to produce an organized, logical sequence of messages, while adjusting the messages to ensure that they are appropriate for the listener.

Narrative discourse production tasks, specifically story narratives, also provide an opportunity to analyze "story grammar" knowledge, defined as the internal structure of stories that guides both comprehension and production of temporal and

causal relationships among people and events (Coelho, 2002). Story grammar components include the *setting*, *initiating event*, *internal response*, *goal*, *attempt*, *outcome*, and *reactions* (Van Den Broek, 1994). The *setting* statements provide the backdrop for the rest of the story by describing characters, objects, geographical information, and temporal information, among other things. *Initiating events* describe events that set the story in motion. An initiating event results in an *internal response*, which is the reaction of the protagonist to the initiating event. The internal response then leads to the establishment of the *goal*. The *attempts* category describes the various ways in which the protagonist tries to reach the goal, which then leads to the *outcome* of the story. The outcome then results in *reactions*, which describe the reaction of the protagonist to the success or failure of the outcome (Van Den Broek, 1994). The events and actions occurring in a story are organized into subplots or *episodes*, which revolve around both a goal and its outcome (Van Den Broek, 1994).

Both linguistic and cognitive processes are thought to be involved in story-telling (Coelho, 2002). This is due, in part, to the complexity of ideas that a story narrative can express, in addition to the logical, structural (not necessarily content-bound) organization of a story grammar episode. In addition, in order to produce narrative discourse, linguistic information must be integrated within an overall theme, or macrostructure. Hence, narrative production tasks also tap the ability to integrate cues underlying the macrostructure (Tucker & Hanlon, 1998).

Many studies have employed tell-retell tasks. While a story tell/retell task is considered to fit within the realm of discourse, its monologue format does not allow for a re-creation of the conventions and subtleties of conversational exchange (Snow,

Douglas, & Ponsford, 1999). In addition, it includes a different set of demands. As such, competence on a story tell/retell task does not imply competence in conversation. However, examining discourse production, especially through narrative production tasks, has been shown to be especially sensitive to subtle language deficits (Tucker & Hanlon, 1998). Ulatowska et al. (1983) note that storytelling is a complex and critical communicative event. In addition, Snow et al. (1999) have suggested that persons who demonstrate difficulty using the narrative genre will have difficulty reconstructing their own life experiences in order to share with others. Clark (1994) suggests that discourse, when viewed as a joint activity, also applies to stories told to others by single narrators. Finally, tell-retell tasks provide a controlled environment. As opposed to more open-ended conversational analyses of discourse, a story narrative task is structured enough to sample the behaviors of interest and yet similar enough to discourse that takes place on a daily basis to elude the difficulties brought about by the use of more artificial tasks.

### **Discourse Impairments in Developmental Language Disorders**

Because the deficits in PFCD are so subtle, it may well be that the learning disabled (LD) population provides a better comparison cohort. There is limited research regarding the discourse of adults with LD. However, available research suggests that they may have difficulty producing coherent and cohesive stories. Also, their written language may present with a lack of overall sense of structure within and between sentences (Gregg & Hoy, 1989). Roth and Spekman (1994) explored oral story production in LD adults compared with non-LD adults. They utilized two story elicitation techniques: picture-elicitation and spontaneous story generation. The

stories of the LD adults conformed to the basic rules of story grammar, demonstrating a fundamental knowledge of narrative schema. However, Roth and Spekman (1994) found that several behaviors differentiated between groups. For example, compared to non-LD adults, LD adults presented with reduced story and episode length, reduced episode completeness, and the use of less sophisticated linguistic markers to connect episodes to one another.

More research has addressed the discourse of student-age children with LD. Findings suggest that they present with some of the same characteristics as the previously discussed RHBD and TBI populations. Specific to expressive discourse, LD students may use shorter utterances, provide less information (McCord & Haynes, 1988; Garret, 1986), demonstrate difficulties with language formulation and organization, produce insufficient story schemas (Montague, Maddux, & Dereshiwsky, 1990), and differ from non-LD students in their cohesive organization and adequacy (Liles, 1985; Garret, 1986). In spontaneously generated stories, Roth and Spekman (1986) found that students with LD showed particular impairment in the story grammar categories of minor setting statements, internal response, attempts, and planning statements, compared to normally achieving (NA) students. Impairment in the internal response category may indicate difficulty interpreting affective or emotional responses, goals, desires, or other internal states of the story protagonists. Roth and Spekman (1986) also found that the stories told by the LD students contained fewer propositions and complete episodes than their NA peers.

## **Methods of Discourse Production Analysis**

Although the discourse production deficits of patients with PFCD can be considered subclinical or mild on the basis of conventional language test batteries, they nonetheless can have profound effects on quality of life, social re-integration, and the overall ability to communicate effectively. Despite the existence of multiple tools capable of examining and quantifying discourse features, none has been designed specifically to be sensitive to capturing the complex deficits of the PFCD population.

Various tools that have been used to analyze discourse production in aphasia include the Clinical Discourse Analysis (Damico, 1985), Linguistic Communication Measure (Menn, Ramsberger, & Helm-Estabrooks, 1994), *Quantitative Analysis of Aphasic Sentence Production* (Rochon, Saffran, Berndt, & Schwartz, 2000; Saffran, Berndt, & Schwartz, 1980), Cohesion Analysis (Mentis & Prutting, 1987), Topic Analysis (Mentis & Prutting, 1991), Intonation Unit Analysis (Wozniak, Coelho, Duffy, & Ziles, 1999), and the application of the Systemic Functional Linguistics approach (Togher, 2001). Some of these tools utilize standard stimuli, such as the “cookie theft picture,” to elicit narrative discourse. None of these tools are computer-assisted, and all are laborious and time-intensive in their use. In addition, many of these tools codify discourse features using general terminology that is difficult to operationally define. (See Appendix A for detailed descriptions of discourse analysis systems).

However, selected characteristics of these tools may be helpful in analyzing the discourse production of patients with PFCD. For example, various parameters of the Clinical Discourse Analysis (Damico, 1985) include assessing insufficient

information bits, message inaccuracy, poor topic maintenance, inappropriate speech style, and inability to structure discourse. The Clinical Discourse Analysis (Damico, 1985) has been applied to various populations, including TBI, but contains no published psychometric evidence of reliability or validity.

The Linguistic Communication Measure (LCM; Menn et al., 1994) quantifies the amount of information conveyed verbally and the proportion of informative to non-informative words produced. The Indices of Lexical Efficiency and Lexical Support could be applied to the PFCD population, who tend to have either enhanced or reduced speech output, by inserting extraneous details or leaving out critical ones, respectively. Saffran et al.'s (1980) *Quantitative Analysis of Aphasic Sentence Production* focuses primarily on syntax, but also measures elaboration, which, similar to the LCM, may be a useful feature of an analysis of discourse production in PFCD patients.

Mentis and Prutting (1987) utilized a system created by Halliday and Hasan (1976) to analyze the discourse cohesion of head-injured and normal adults, both in narrative and conversational formats. Six cohesion categories were analyzed, including lexical, reference, ellipsis (the omission of a word or words that are understood but that must be supplied to make a construction grammatically complete), conjunction, substitution, and incomplete cohesion, of which reference, ellipsis, conjunction, lexical, and incomplete cohesion may be applicable to the PFCD population.

Mentis and Prutting (1991) created a multidimensional topic analysis instrument designed to be sensitive to problems and patterns in topic management in



head-injured and normal adults. They applied the analysis to both monologue and conversational formats. In the monologue format, comprehensiveness of topic was analyzed, along with topic and subtopic maintenance. Both parameters may be useful when analyzing the discourse production of the PFCD population.

Wozniak et al. (1999) employed a modified version of the intonation unit analysis developed by Mentis & Prutting (1991). Wozniak et al. (1999) segmented conversational samples into intonation units, and then placed each intonation unit into an ideational intonation category. Ideational intonation categories included: contains new information, no new information, incomplete, or tangential. The “tangential” category is the only one that may be helpful in analyzing the discourse production of the PFCD population, but is so broadly defined that it is difficult to operationalize.

In contrast to these methods for coding output from aphasic patients, two computer-assisted systems, namely Computerized Language Analysis (CLAN; MacWhinney, 2000) and Systematic Analysis of Language Transcripts (SALT; Miller & Chapman, 1985) have been created to analyze language transcripts systematically and quantitatively. Although they have been used primarily to analyze child language, they can also be adapted to analyze adult language. In addition, although existing discourse codes are oriented towards analysis of interactive discourse, the specific codes created for CLAN can be manipulated, and new codes added, thus enhancing the flexibility of CLAN as a tool for discourse analysis. Unlike CLAN, SALT is limited in its flexibility, allowing the use of existing codes only.

CLAN is designed to analyze data that is transcribed in “CHAT” format, and consists of a graphic interface. It completes analyses via a series of commands that search for strings and compute various linguistic indices. In general, CLAN enables various programmed analyses of transcribed data, including frequency counts, word searches, co-occurrence analyses, and morphosyntactic analysis.

The CLAN programs that can be helpful in analyzing the basic features of narrative transcripts include mean length of utterance (MLU), type-token ratio (TTR), and frequency counts of researcher-directed features. The “FREQ” (“frequency”) program performs frequency counts that involve calculating the number of times a word or specified feature occurs in a file or set of files. This program produces a set of all words in the specified file and their frequencies, along with a type-token ratio (TTR). The “MLU” program calculates the MLU of an entire file or a specified subset of a file. MLUw (words) can also be calculated, which doesn’t require identifying the morphemes in words.

Overall, CLAN allows for a sufficient level of flexibility to pilot discourse codes. The current study will utilize CLAN codes that are already in existence (listed above) as well as novel codes to tap the unique discourse deficits of patients with PFCD. The novel codes will be based on a discourse error analysis. The categories of error analysis, along with their operational definitions, are listed in Appendix B.

## **Rationale for the current study**

Because general descriptions of linguistic deficits yielded from conventional language tests do not adequately characterize the discourse deficits of PFCD patients, discourse analysis may be the most appropriate and sufficiently sensitive method to

effectively and quantitatively analyze the language deficits of PFCD patients. No set of codes exists currently to describe the unique discourse production deficits of the PFCD population. This proposal outlines the development of codes designed to differentiate between the “nonaphasic” but characteristic discourse production deficits that could present following PFCD and the discourse of a healthy comparison group.

Specifically, this preliminary study was designed to answer the following research questions:

1. Do the coding categories demonstrate sufficient inter-rater reliability?
2. Do the scores and indices derived from the discourse analysis differ significantly between the PFCD group and matched healthy comparison group members on the indices of content units, thematic units, story grammar features, and discourse errors?
3. Do the scores and indices derived from four levels of analysis (content units, thematic units, story grammar components, and discourse error analysis) differ significantly between the participants tell and retell scores by group, with each group performing more poorly on the retell?
4. Does a qualitative error analysis profile differentiate between the two study groups, with the PFCD population demonstrating discourse errors classically associated with PFCD, and the comparison group exhibiting relatively few such errors?

## **Method Overview**

The pilot study consisted of two phases. In Phase I, experimental stimuli were modified from a pictorial story, *Old MacDonald Had an Apartment House* (Barrett &

Barrett, 1969; Gernsbacher & Varner, 1988), and piloted with healthy volunteers. Participants viewed the entire computerized pictorial story and then told the story with as much detail as possible. The story-tell was audiotaped for future transcription. Participants then answered 15 multiple-choice comprehension questions relating to story content (adapted from Gernsbacher & Varner, 1988). Questions tapped inferencing, factual details, time concepts, gist, and cause/effect relationships. In Phase II, PFCD and matched comparison participants viewed the 16-frame computerized pictorial story and then told the story with as much detail as possible (see Appendix E). Participants then retold the story after a 30-minute delay, after which they answered 13 multiple-choice questions relating to story content. Both the story tell and retell were audiotaped for future transcription. The narrative codes were developed, based on both the literature and pilot data, and were then applied to the transcripts of patients with PFCD and healthy comparison group members.

## **PHASE I**

### **Purpose**

The purpose of the pilot was to test the appropriateness of the task in terms of its level of difficulty, clarity of comprehension questions, appropriateness of time limits, and the participants' ability to generate adequate story narratives, in terms of both frame-by-frame descriptions and overall gist. Pilot results were used to adapt and modify the instructions, pictorial frames, and comprehension questions accordingly.

## Piloting of the Narrative Discourse Task

The preliminary story-tell task (19 frames) and original fifteen comprehension questions were piloted with healthy volunteers (N=14). Pilot participants ranged in age from 23 to 68 years (mean = 36; SD  $\pm$  16.9). For the purpose of analysis, the pilot participants were split into two age groups: older and younger. The younger participant group (N=9) ranged in age from 23 to 34 years (mean = 26; SD  $\pm$  4), while the older participant group (N=5) ranged in age from 58 to 69 years (mean = 65; SD  $\pm$  4). For additional demographic information, see Table 1.

Table 1. Number and percentage of participants according to age group by gender, education level, and race/ethnicity

Characteristic	Younger Group		Older Group	
	Number	Percentage	Number	Percentage
Gender				
Female	8	89	3	60
Male	1	11	2	40
Education Level				
Some College	0	0	1	20
College	8	89	2	40
Masters	1	11	1	20
PhD	0	0	1	20
Race/Ethnicity				
White, non-Hispanic	8	89	4	80
Black, non-Hispanic	1	11	1	20

NOTE: Percentages may not sum to 100% due to rounding.

All sessions were audiotaped using a Sony TCM-20DV Cassette-Corder and a Labtec Verse 504 external microphone. The task was presented on a Macintosh G4 laptop in a room at the University of Maryland that was free of auditory and visual distractions.

## Specific Task Details

Participants were verbally presented with a standard set of instructions, which were also simultaneously presented on the computer screen. The instructions described the task and what responses were expected of the participant. Each participant was told,

“This task helps me understand how people produce stories. You will see a series of pictures that tell a story. After viewing and understanding each picture, press the SPACE bar to continue. If you have not responded within 20 seconds, the computer will move to the next picture. Some of the frames will be detailed, so take your time and look carefully.” Participants were then told, “After you have finished viewing the pictures, I want you to, ‘**Tell me the story, from the beginning to the end.**’ Your telling of the story will be audio tape-recorded. At that time, you will be asked to respond to a set of questions about the story. So I want you to remember as much as you can about it.”

Subsequent to the instructions, participants were presented with a title page, which included both the story title and reference. Each pictorial frame of the story was presented for a maximum of 20,000 ms. Participants controlled the pace of presentation, within the 20,000 ms limit, by pushing the space bar when they were ready to move on to the next frame. If the space bar had not been pushed within 20,000 ms, the computer was programmed to move on to the next frame. Response times for processing each frame were measured in milliseconds.

Following the presentation of the story, a prompt appeared on the screen, instructing the participant to “Tell the story now, from beginning to end. Be as specific as possible.” Following the story-tell, the subject then pressed the space bar

to reach the following prompt, “You will now respond to a set of multiple-choice questions, and the responses will be A, B, C, or D. The questions will appear on the computer screen, one at a time. To select your answer, use your index finger to press the labeled key corresponding to your answer.” The A, B, C, or D responses were labeled on the number pad of the keyboard. Next, participants were told, “Both accuracy and speed are important. I want you to respond quickly to the questions, but not so quickly that you make mistakes. If you have not responded within 10 seconds, the computer will go on to the next question. If you are not sure of an answer, go ahead and guess.” The presentation of the comprehension questions was similar to that of the pictorial frame presentation, in that response times were measured and participants could control the pace of presentation within a maximum time of 10,000 ms. If the participant did not respond within 10,000 ms, the computer was programmed to move on to the next question.

The pilot version of the task also included a 7-item questionnaire aimed at providing the investigators with additional information about perceived level of task difficulty, timing features, and ability of the comprehension questions to tap both implicit and explicit information presented in the story (See Appendix C).

Each narrative generated from Phase I was reviewed three times by primary author, IE. Prior to analysis, the salient information conveyed by each pictorial frame was described by IE and co-author CF. Subsequent to each of the three reviews, these frame-by-frame content descriptions were compared with the appropriate components of each individual’s story-tells, in order to objectively analyze the content. Content for each frame was then scored as present or absent. In addition,

four thematic units, covering the range of thematic content/gist of the story, were identified by the same two investigators and then compared with the individual's story-tells. These were scored as either present or absent (e.g., 1 or 0), or given partial credit (.5).

## Results

Frame-by-frame analysis revealed that while one frame was not described by any of the participants, the remaining frames were described by between 33% (N=3) and 100% (N=9) of the younger participants and between 20% (N=1) and 100% (N=5) of the older participants (See Table 2).

Table 2. Number and percentage of participants who referenced frame content in story tell according to age by frame number and description.

Frame number and description	Younger Group		Older Group	
	Number	Percentage	Number	Percentage
1 Old MacD and wife in front of apt. bldg., they move to the city.	9	100	4	80
2 Old MacD is sup't of bldg. Wife is sad and holding dying plant that is not getting enough light.	6	66	2	40
3 Old MacD cuts down bushes in front of window to let the sunshine in.	3	33	2	40
4 Plant now has light and is thriving.	4	44	1	20
5 Old MacD begins planting seeds outside the building. Tenants are upset.	4	44	1	20
6 The first tenants move out because they are unhappy with the changes.	6	66	2	40
7 Old MacD begins growing vegetables inside the apt. bldg.	9	100	5	100
8 Carrots grow through ceiling of one apt., apartment dwellers are angry.	7	77	2	40
9 One couple hear a cow through the wall and are upset. The man has a rifle.	8	88	3	60
10 Fruits and vegetables are growing inside the apts. More tenants are forced to move out.	8	88	4	80
11 Bldg. owner furious when sees that apt. bldg. has been converted into a "farm."	8	88	2	40
12 Bldg. owner gets angry at Old MacD for converting the apt. bldg.	6	66	2	40
13 Sad Old MacD and wife packing to leave.	5	55	2	40
14 Owner tries to decide what action to take because the plants are thriving.	6	66	1	20
15 Owner happy, standing in front of construction of a fruit and vegetable stand. Old MacD & wife shocked.	7	77	2	40
16 Owner with Old MacD and wife in fruit and vegetable stand; many customers.	8	88	4	80
17 Winter scene; apartment "farm" and fruit and vegetable stand thriving.	0	0	0	0



Each of the four thematic units (See Table 3) was described correctly by >60% of the younger participants. The first and last thematic units were described by 100% and 88% of the younger participants, respectively, while the middle two were each described by 61% of the participants. The first and last thematic units were described by 100% and 60% of the older participants, respectively, while the middle two were described by 50% and 40% accuracy, respectively. It should be noted that the first and last thematic unit contained the most salient thematic aspects of the story, while the middle two represented the story's build-up or progression.

Table 3. Number and percentage of participants who referenced thematic content unit according to age by thematic unit number and description.

Thematic unit number and description	Younger Group		Older Group	
	Number	Percentage	Number	Percentage
1 Love of farming spurs Old MacD, the sup't of an apt. bldg., to start growing produce in the bldg.	9	100	5	100
2 As the apt. "farm" thrives, Old MacD forces the tenants to leave to accommodate farm.	5.5	61	2.5	50
3 The bldg. owner is angry and evicts Old MacD and his wife.	5.5	61	2	40
4 The bldg. owner thinks of a "win-win" situation; let the Old MacD's stay and open a fruit and vegetable stand.	8	88	3	60

An analysis of responses to the comprehension questions revealed that, for those who responded, the percentages correct per question ranged from 29% to 100% in the younger participant group (See Table 4). In the older participant group, the percentages correct per question of respondees ranged from 25% to 100% .

Table 4. Percentage of correct responses according to age by question number.

Question #	Younger Group		Older Group	
	% Correct	% No Response	% Correct	% No Response
1	29	22	66	40
2	38	11	50	20
3	50	11	100	0
4	75	11	100	0
5	78	0	25	20
6	75	11	60	0
7	67	33	66	40
8	50	11	80	0
9	67	0	66	40
10	50	11	40	0
11	100	33	66	40
12	86	22	80	0
13	75	11	50	20
14	100	0	100	20
15	89	0	80	0

The results of the pilot questionnaire (See Table 5) indicated that 66% (N=6) of younger participants found the story easy to moderately easy to understand, 33% (N=3) found it difficult to understand, and 0% (N=0) found it very difficult to understand. Seventy-seven percent of younger participants (N=7) reported difficulty processing the pictures. In terms of timing, while 77% (N=7) of the younger participants reported having sufficient time to process each pictorial frame, only 11% (N=1) reported having sufficient time to respond to the comprehension questions (set at a maximum of 10 seconds). All younger participants reported that the comprehension questions were clearly worded and 66% (N=6) felt that the comprehension questions fairly tapped the information presented in the story.

The results of the pilot questionnaire indicated that 80% (N=4) of older participants found the story easy to moderately easy to understand, 0% (N=0) found it difficult to understand, and 20% (N=1) found it very difficult to understand. Sixty

percent of older participants (N=3) reported no difficulty in processing the pictures. In terms of timing, 80% (N=4) of the older participants reported having sufficient time to process each pictorial frame and sufficient time to respond to the comprehension questions (originally set at a maximum of 10 seconds). All older participants reported that the comprehension questions were clearly worded and fairly tapped the information presented in the story.

Table 5. Number and percentage of participants according to age by responses to questionnaire.

<b>Question responses</b>	<b>Younger Group</b>		<b>Older Group</b>	
	<b>Number</b>	<b>Percentage</b>	<b>Number</b>	<b>Percentage</b>
Ease of understanding story				
Easy	2	22	2	40
Moderately easy	4	44	2	40
Difficult	3	33	0	0
Very difficult	0	0	1	20
Any pictures difficult to process				
Yes	7	77	2	40
No	2	22	3	60
Enough time to process each frame				
Yes	7	77	4	80
No	2	22	1	20
Questions clearly worded				
Yes	9	100	5	100
No	0	0	0	0
Questions fair				
Yes	6	66	5	100
No	3	33	0	0
Enough time to answer each question				
Yes	1	11	4	80
No	8	88	1	20

NOTE: Percentages may not sum to 100% due to rounding.

The pilot questionnaire also included open-ended questions. Responses to this portion of the questionnaire indicated that some pictures towards the beginning of the story were overly detailed and somewhat difficult to process, that the comprehension questions were too specific and difficult to catch when reading for the “gist,” and that despite some confusion at the beginning of the story, the story became easier to understand as it progressed (see Appendix D).

### **Task Revisions, Based on Pilot Results**

The pilot task underwent multiple revisions based on the pilot results. The pilot task gave participants 10,000 ms to answer the comprehension questions before moving on to the next frame. Because between 0% to 40% of the participants did not answer any given question, the maximum amount of time provided to answer the comprehension questions was extended to 30,000 ms. The four questions that were answered correctly by less than 50% of the respondents were either replaced or modified, and the wording of four additional questions was also revised, leaving 13 remaining questions. Other adjustments included adapting the instructions by adding a frame that notifies the participant of upcoming comprehension questions, and splitting two of the story frames into 4 separate frames in order to decrease the amount of detail presented in a single pictorial frame. In addition, four frames were enlarged in order to clarify the images and make the details easier to see.

Several frames not critical to story coherence and gist were eliminated from the original story, leaving sixteen frames (see Appendix E). The results of the pilot suggested that several of these non-critical frames were confusing to the participants. For example, one frame depicted a man and his wife tearfully standing in the hallway

of the apartment building, carrying their luggage. The couple is presumably being kicked out of the apartment complex by Old MacDonald's burgeoning farm. The man in this frame was often confused with Old MacDonald, which led to overall confusion about the story line. This frame was not critical to the story, and as such, was removed. Finally, one of the thematic units was split into two, to eliminate the need for partial credit.

The pilot participant group was quite heterogeneous demographically, so much so that in order to provide a reasonable analysis, the group was split into two groups (older and younger). The groups differed in demographics such as education level as well; while 89% of the younger group had obtained a college education, only 40% of the older group had done so. As such, every attempt was made to closely match participants in Phase II of the study.

Finally, as Phase I was solely intended to pilot the story task itself along with preliminary codes and comprehension questions, participants were not asked to retell the story after a 30-minute delay. This portion of the task was added in Phase II.

## **PHASE II**

### **Participants**

Ten participants were included in Phase II, the preliminary descriptive study: five with left, right, or bi-frontal lesions confined to the PFC (N=5; 3 males; Mean age=54 [SD±12.6]; Mean education=17.8 [SD±2.68]), and five age-, gender- and education- matched non-brain damaged healthy comparison group members (N=5; 3 males; Mean age=52 [SD±13.8]; Mean education=17 [SD±2.55]) (See Table 6). Participants ranged in age from 38 to 73 years. For PFCD patients, CT or MRI scan

reports, as interpreted by a neuroradiologist blinded to medical diagnosis, confirmed focal prefrontal cortex lesions (either left-, right- or bilateral dorsolateral [Brodmann's areas 8,9, 46], orbital [Brodmann's areas 10-13, 47], or medial/cingulate [Brodmann's areas 8-10, 23, 34, 31]).

Table 6. Demographic Data, Reaction Times, and Comprehension Question Accuracy, by Group

Group	Age (years) (M, SD, Range)	Years of Education (M, SD, Range)	Gender	ProcessingTime (M, SD)	Comp Question Accuracy (M, SD, Range)
PFCd	54 (12.6) [40-72]	17.8 (2.68) [15-22]	2F; 3M	458.97 (146.51)	10 (2.345) [7-12]
Controls	52 (13.8) [38-70]	17 (2.55) [13-19]	2F, 3M	474.98 (43.3)	11.6 (1.140) [10-13]
P-levels	0.85, ns	0.723, ns		0.519, ns	0.256, ns

Lesion location (in terms of Brodmann Area Intersection) and volume were determined from MR and CT images using the Analysis of Brain Lesion (ABLE) software (Makale, Solomon, Patronas, Danek, Butman, & Grafman, 2002) contained in MEDx v3.422 (Sensor Systems, Inc., Sterling VA) with enhancements to support the Volume Occupancy Talairach Label (VOTL) database (Lancaster, Woldorff, Parsons, Liotti, Freitas, Rainey, Kochunov, Nickerson, Mikiten, & Fox, 2000). As part of this process, the MR or CT image of the brain was spatially normalized to Talairach space (Talairach & Tournoux, 1988). Intersection of Brodmann areas was determined by automatic queries to the VOTL database. Lesion volume was determined by manual tracing of the lesion in all relevant slices of the MR/CT image, then summing their areas and multiplying by slice thickness (Table 7).

Table 7. Mean Lesion Size, by Group

Group	Mean Lesion Volume	Range of Lesion Size
Left PFCD (N=1)	48.5	N/A
Right PFCD (N=3; data missing on 1 participant*)	28.6	24.4-32.8
Bilateral PFCD (N=1)	96.3	N/A

\* A CT scan was not performed on one participant with right PFCD. As such, mean lesion volume and size were not computed.

Three participants presented with lesions confined to the right PFC, one with a lesion confined to the left PFC, and one with bilateral PFC damage. All five PFCD group participants were fairly mild in their lesion size/volume, and as such, it was expected that their presentations be fairly mild on the PFCD spectrum. Time post-onset of PFCD ranged from 1 to 12 years. Participants were pair-matched for age, education, and gender. In the one-to-one matching, differences in age ranged from .5 to 1.75 years, while differences in years of education ranged from 0 to 4.

All participants were right-handed, native English speakers. All participants demonstrated adequate visual acuity to read 15-point print and adequate manual dexterity to press keyboard keys in response to stimuli. All participants also passed a visual scanning and tracking screening with  $\leq 1$  error on a total of 38 targets. All participants were administered a battery of assessments, including: *The Discourse Comprehension Test* (Brookshire & Nicholas, 1993), the *Auditory Comprehension Test* (Shewan, 1979), and the *Dementia Rating Scale-2* (Jurica, Leitten, & Mattis, 2001). The PFCD patients were also administered the *Wechsler Memory Scale – 3* (Wechsler, 1997). The *Discourse Comprehension Test* is a receptive assessment that

provides information about a reader's comprehension and retention of four types of information from a homogeneous set of stories. On the *Discourse Comprehension Test*, the mean score for the PFCD group was 36 (SD 1.58), while the mean score for the comparison group was 37.8 (SD .84). The *Discourse Comprehension Test* scores differed significantly ( $t(4)=-3.67$ ,  $p = .021$ ). The *Auditory Comprehension Test* is a receptive assessment in which the participant points to one of four pictures that best illustrates the meaning of a sentence spoken by the clinician. The mean *Auditory Comprehension Test* score for both groups was 20 (PFCD SD = 1.22; Comparison Group SD = 1) ( $t(4).000$ ,  $p = 1.00$ , ns). The *Dementia Rating Scale – 2* is intended to measure and track mental status in adults with cognitive impairment. It assesses both receptive and expressive skills across a variety of subtests, including: attention; initiation/perseveration; construction; conceptualization; and memory. The mean *Dementia Rating Scale – 2* score for the PFCD group was 140 (SD 2.35), while the mean score for the comparison group was 142.8 (SD 2.17) ( $t(4)=-1.532$ ,  $p = .20$ , ns). The *Dementia Rating Scale – 2* also yields an age-corrected MOANS scaled score (AMSS), which reflects where the individual's score lies compared to the distribution of total scores in the normative age group. On the AMSS yielded from the *Dementia Rating Scale – 2*, the mean PFCD group score was 10.6 (SD 2.3), while the mean comparison group score was 15.4 (SD 4.28) ( $t(4)=-1.79$ ,  $p = .147$ , ns). No statistically significant differences in assessment scores were noted between the PFCD group and healthy comparison group members with the exception of the *Discourse Comprehension Test*. It is interesting to note that, consistent with the rationale for this study, all but one of the tests administered failed to distinguish



between groups. This supports the notion that traditional assessments are not sensitive to the deficits found subsequent to PFCD.

PFCD patients were also administered the *Wechsler Memory Scale – 3* (Wechsler, 1997). The *Wechsler Memory Scale – 3* provides subtest and composite scores that assess memory and attention functions using both auditory and visual stimuli. One participant's performance on the *Primary Scaled Indexes* of the *Wechsler Memory Scale – 3* was consistently lower than the others, ranging from almost two SD below the mean to the assessment mean of 100. This participant's scores would be considered clinically significant, and indicate weaknesses in various memory skills, including Immediate Memory, General (Delayed) Memory, and Visual Immediate Memory. Other participants' performances on Primary Scaled Indexes ranged from 1 SD below the mean to 1.5 SD above the mean and would not be considered clinically significant. However, it is important to note that even within our small sample size, the participants were not normally distributed. For further information on subject test profiles, see Table 8.

Healthy comparison participants had no history of neurological or psychological disorders, developmental dyslexia, or substance abuse and passed the *Dementia Rating Scale* (Jurica, Leitten, & Mattis, 2001) with a cutoff AMSS (MOANS scaled score) of 5.

Table 8. Test Score Data, by Group

Group	ACT (M, SD, Range)	DCT (M, SD, Range)	DRS-2 (M, SD, Range)	DRS-2 AMSS (M, SD, Range)
PFC	20, 1.22, 18-21	36, 1.58, 34-38	140, 2.35, 138-143	10.6, 2.3, 9-14
Controls	20, 1.00, 19-21	37.8, 0.84, 37-39	142.8, 2.17, 139-144	15.4, 4.28, 9-19
P-levels	1.00, ns	0.021	0.2, ns	0.147, ns

Participants were recruited at the National Institutes of Health (NIH, W.G. Magnuson Clinical Center). Recruitment took place under a research protocol conducted at the W.G. Magnuson Clinical Center, Rehabilitation Medicine Department (Protocol # 00-CC-0096: Investigations in Discourse Processing, Carol Frattali, PI).

### **Experimental Stimuli**

As described above, experimental stimuli were modified from the pictorial story, *Old MacDonald had an Apartment House* (Barrett & Barrett, 1969), which is also included in Gernsbacher and Varner's (1988) *Multi-media Comprehension Battery*. The final story task consisted of 16 pictorial frames and 13 multiple-choice comprehension questions related to story content (adapted from Gernsbacher & Varner, 1988). The questions were designed to tap inferencing, factual details, time concepts, gist, and cause/effect relationships.

Pictorial stimuli and instructions for the experimental task, along with comprehension questions, are provided in Appendix E.

## **Procedure**

Participants were seated directly in front of a Macintosh G4 computer with an 18-inch View Sonic monitor in a test environment free from auditory and visual distractions. Story tells and retells were audiotaped by a Panasonic Standard Cassette Transcriber, Model No. RR-830. Participants were asked to view the pictorial story, displayed frame-by-frame on the computer, programmed in SuperlabPro™ 1.74. Once they finished viewing the entire story, participants were asked to tell the story with as much detail as possible. After a thirty-minute delay, participants were asked to re-tell the story, subsequent to which they answered a series of multiple-choice comprehension questions. Because several changes were made to the task subsequent to the pilot study, specific task details will be presented again, although some of these details were presented under Phase I *Task Details*.

## **Specific Task Details**

Participants were verbally presented with a standard set of instructions, which were also simultaneously presented on the computer screen. The instructions described the task and what responses were expected of the participant. Each participant was told,

“This task helps me understand how people produce stories. You will see a series of pictures that tell a story. After viewing and understanding each picture, press the SPACE bar to continue. If you have not responded within 20 seconds, the computer will move to the next picture. Some of the frames

will be detailed, so take your time and look carefully.” Participants were then told, “After you have finished viewing the pictures, I want you to, **‘Tell me the story, from the beginning to the end.’** Your telling of the story will be audio tape-recorded. Later on, I will ask you again to tell me about this story. At that time, you will be asked to respond to a set of questions about the story. So I want you to remember as much as you can about it.”

Subsequent to the instructions, participants were presented with a title page, which included both the story title and reference. Each pictorial frame of the story was presented for a maximum of 20,000 ms. Participants controlled the pace of presentation, within the 20,000 ms limit, by pushing the space bar when they were ready to move on to the next frame. If the space bar had not been pushed within 20,000 ms, the computer was programmed to move on to the next frame. Response times for processing each frame were measured in milliseconds.

Following the presentation of the story, a prompt appeared on the screen, instructing the participant to “Tell the story now, from beginning to end. Be as specific as possible.” Following the story-tell, the participant was engaged in other protocol activities for half an hour, such as the administration of portions of the assessment battery that was described under the Phase II *Participants* section. After 30 minutes, the participant was told, “Do you remember the story you viewed earlier? I want you to tell me the story now, from beginning to end. Be as specific as possible.” Following the retell, the subject then pressed the space bar to reach the following prompt, “You will now respond to a set of questions about the story. To get to the questions, press the space bar. You will respond to a set of multiple-choice

questions, and the responses will be A, B, C, or D. The questions will appear on the computer screen, one at a time. To select your answer, use your index finger to press the labeled key corresponding to your answer.” The A, B, C, or D responses were labeled on the number pad of the keyboard. Next, participants were told, “Both accuracy and speed are important. I want you to respond quickly to the questions, but not so quickly that you make mistakes. If you have not responded within 30 seconds, the computer will go on to the next question. If you are not sure of an answer, go ahead and guess.” Finally, the participants were presented with a screen stating, “Questions: Press any key to answer the questions.” The presentation of the comprehension questions was similar to that of the pictorial frame presentation, in that reaction times were measured and participants could control the pace of presentation within a maximum time of 30,000 ms. If the participant did not respond within 30,000 ms, the computer was programmed to move on to the next question.

## **Data Analysis and Reduction**

### **Transcription of Sample/Coding**

The first author, blinded to group assignment, transcribed each transcript orthographically, utterance by utterance, into the Computerized Language Analysis system (CLAN; MacWhinney, 2000), a system used to systematically and quantitatively code and analyze language transcripts. Each transcript was then analyzed on four levels: content units, thematic units, story grammar categories, and discourse errors.

Three levels of coding (content units, thematic units, and story grammar categories) were based on a template of the story created by authors IE and CF. First,

a template of 16 frame-by-frame content descriptions was created. Each transcript was reviewed utterance by utterance, and then compared to the frame-by-frame descriptions in order to objectively analyze story content. Each frame of the story was assessed as being adequately or inadequately covered in the story-tell and retell, and then scored as present or absent (e.g. 1 for present, 0 for absent). The actual scoring did not take place within the transcript, but rather in a separate spreadsheet that was laid out according to frame number (See Appendix F). In addition, 5 thematic units, covering the range of thematic content/gist of the story, were identified by the same two investigators and then compared with the individual's story-tells. Again, these were scored as either present (1 point) or absent (0 points) in a separate spreadsheet (See Appendix F). For the story grammar analysis, one story grammar category was identified for each frame of the frame-by-frame description (with the exception of one frame which contained two story grammar categories). The story grammar categories consisted of setting, initiating event, internal response, goal, attempt, outcome, and reactions (Van Den Broek, 1994). These 17 frame-by-frame story grammar associations were compared with each transcript and scored as present or absent in a separate spreadsheet. Story grammar analysis differed from content unit analysis in that each frame needed to be consistent with the intent and structure of the story grammar category in order to be scored as present (See Appendix F). Overall, participants could obtain a maximum total score of 16 content units, 5 thematic units, and 17 story grammar units for each story-tell or retell.

The fourth level of analysis (see Appendix B) consisted of a computer-assisted discourse error analysis. Twenty codes were identified to tap narrative

discourse deficit features typically associated with PFCD. These features were organized into three overall categories: word level errors, phrase/sentence level errors, and global/gist errors. Error codes created at the word level included: superordinate substitutions, word-finding difficulties such as phonemic/formal/semantic/unrelated paraphasias and unretrieved words, perseveration, and violation of cohesive links (anaphora and conjunction). Error codes created at the phrase level included: embellishment, confabulation, perseveration, topic stray, faulty inference (backward, predictive, coherence), and faulty temporal ordering of events. Errors at the thematic/global level included difficulty interpreting gist, which was defined as partial or narrow ability or inability to capture the theme or gist of the overall story (Frattali & Grafman, in press). Appendix B provides a comprehensive listing of each error type, the manner in which it was coded within the transcript, the error category (e.g. word level, phrase/sentence level, and thematic/global level), an operational definition, and an example of each error. All of these features were coded via an error analysis in which their presence was noted within the transcript. More specifically, when an error occurred in an utterance, codes identifying that error were entered in the line beneath the CHAT main tier utterance, also known as the “dependent tier.” Each error was first identified by its level, e.g., word, phrase/sentence, or global. If multiple errors occurred on the same level (e.g., word) in an utterance, they were all listed in one dependent tier. If multiple errors occurred in an utterance, but some were at the word level and some were at the phrase level, two different dependent tiers were created beneath the utterance, and errors were categorized according to level. For each story

tell and retell, errors in each discourse error category were summed, along with total errors at the word, phrase, and global levels.

### **Selection of Final Codes for Data Analysis**

Once the audiotaped samples were transcribed verbatim, the primary judge (first author IE) coded all samples on the four levels of analysis, as described above. Two additional independent volunteer judges blinded to subject and group assignment were trained to code the transcripts for the purpose of examining inter-rater reliability. Training took place over a one and one-half hour time period in a quiet space free from distractions. Judges were provided with a comprehensive training manual (See Appendix F), and reviewed the entire manual with the primary judge. The training manual included a summary of the project and detailed procedural information, including the order in which to code, instructions regarding coding at each of the four levels of analysis, and a description of story grammar and story grammar categories. A detailed description of the discourse errors contained in the error analysis was also included, in addition to examples of each error type, and the method to code it within the transcript (See Appendix B). Finally, the manual included a checklist to help the judges keep track of their work. Judges were provided with a random sample of 30% of total transcripts (3/10 total transcripts; each transcript contained both a story tell and retell) to assess inter-rater reliability. The order of tell and retell presentation was also randomized. Final data and codes for content units, thematic units, and story grammar features were assigned on a majority rules basis. Due to the solid reliability of the discourse error analysis, final



data for the discourse errors were determined by the primary judge. The threshold for significance was set at either  $p < .025$  or  $p < .05$ .

## **Data Analysis**

Data analysis procedures were intended to address the following questions: Do the coding categories demonstrate sufficient inter-rater reliability? Do the scores derived from the discourse analysis differentiate between the PFCD group and matched healthy comparison group members on the indices of content units, thematic units, story grammar features, and discourse errors? Do the scores and indices derived from four levels of analysis (content units, thematic units, story grammar components, and discourse error analysis) differentiate between the participants tell and retell scores by group, with each group performing more poorly on the retell? And finally, does a qualitative error analysis profile differentiate between the two study groups, with the PFCD population demonstrating discourse errors classically associated with PFCD, and the comparison group exhibiting relatively few such errors?

## **Results**

The results section will be organized in the following manner. First, inter-rater reliability will be discussed. Next, a post-hoc analysis will be discussed. Then, descriptive statistics will be presented for a general overview of the answers to the remaining research questions. Finally, results will be explained within the context of each of the three remaining research questions, followed by a summary of data trends.

## **Do the Discourse Codes Demonstrate Sufficient Inter-rater Reliability?**

A Kappa statistic for categorical data was used as a measure of inter-rater agreement for three of the levels of analysis: content units, thematic units, and story grammar categories. Kappa is a measure of inter-rater agreement that examines whether each raters counts differ from what would be expected by chance. A Kappa statistic was determined by comparing the codes assigned by each judge to those assigned by the other two judges (e.g., comparing the codes of judge 1 to judge 2, and judge 1 to judge 3, etc.) According to a guideline for interpretation by Fleiss (1981, p. 218), a Kappa value above .75 denotes excellent agreement, values between .40 and .75 denote fair to good agreement, and values below .40 denote poor agreement. Inter-rater reliability was found to be primarily within the fair to good range (See Table 9), with the exception of the poor thematic unit agreement between judges 1 and 3 (.321). Overall, the weakest agreements occurred between judges 1 and 3. Descriptively, judge 3 tended to code far fewer items than judges 1 and 2. Within judge pairs, agreement for story grammar units and content units tended to be more consistent than thematic units. A potential explanation is that only five possible thematic units were coded as present or absent in any given story tell or retell, as opposed to much larger numbers of content units and story grammar components (16 and 17, respectively). In addition, judging the presence or absence of thematic units was more subjective than coding content units and story grammar components. The subjectivity involved in coding thematic units was due to its nature as a measure more consistent with overall gist than frame-by-frame descriptions.

Table 9. Kappa Statistic for Inter-Rater Reliability

Category	Judges 1/2	Judges 2/3	Judges 1/3
Content Units	0.672	0.509	0.48
Thematic Units	0.494	0.737	0.321
Story Grammar Units	0.633	0.553	0.489

As opposed to binary data, the discourse error analysis yielded continuous data. As such, the Kappa statistic was no longer appropriate to assess inter-rater reliability. Instead, an Intraclass Correlation Coefficient (ICC), was utilized to yield a single index to describe reliability (Portney & Watkins, 2000). While six equations can be used to calculate the ICC, Model 2, which assumes that the discourse error analysis could be used by all equally trained clinicians, was used for this study. According to a guideline for interpretation by Portney and Watkins (2000), ICC values above .75 suggest good reliability, and those below .75 denote poor to moderate reliability. However, Portney and Watkins (2000) note that these are not absolute standards, and that judgments should be made within the context of the individual study. Finally, due to the quantity of categories analyzed in the discourse error analysis (N=20), and the relatively small n (n=10), the ICC analysis was restricted to overall measures, including counts of total discourse errors at the word level (by tell and retell), total discourse errors at the phrase level (by tell and retell), and difficulty interpreting gist (by tell and retell).

The ICC for the discourse error analysis ranged from poor to excellent (See Table 10). Word and phrase level analyses ranged from good to excellent. Poor ICC scores for the gist measure suggest very little agreement between judges as to the presence or absence of adequate gist in the story tells/retells.

Table 10. Intraclass Correlation Coefficient – Total Discourse Errors

Word Level – Tell	0.85778
Word Level – Retell	0.65982
Phrase Level – Tell	0.60455
Phrase Level – Retell	0.904
Gist – Tell	0.2
Gist – Retell	0.2

### **Post-hoc Analysis Procedure: Lexical Efficiency**

During the transcription process, it was noted that the PFCD group appeared to tell longer stories and have more difficulty expressing story content than the healthy comparison group. A lexical efficiency measure, defined as the ratio of content units or ideas to the total words produced (Menn et al., 1994), has been utilized in the past to evaluate the narratives of patients with aphasia (Menn et al., 1994). Although calculating lexical efficiency was not an initially posited research question, it was calculated to determine the difference in the relative fullness or emptiness of speech between groups. The procedure for counting the total number of words was completed as described by Menn et al. (1994), with a few modifications. Menn et al. (1994) suggest excluding fragments if they are identifiable as false starts on a word that is eventually produced. We chose to include those false starts, reasoning that these production events contribute to the emptiness of speech that the lexical efficiency measure is intended to capture. In addition, instead of identifying the number of content units as they are traditionally defined, we chose to use content units as they are defined within the context of this study (e.g., yielded from frame-by-frame descriptions), to serve as the numerator of the ratio.

More specifically, lexical efficiency was calculated as follows. First, the total number of words was identified by counting all words in each tell and retell

transcript, including incorrect words, paraphasias, jargon, repetitions, self-corrections, irrelevant statements, digressions, comments, word fragments (Menn et al., 1994), and false starts. Fillers such as *um*, *er*, *uh*, *huh*, and *hmm* were excluded from the total number of words. Next, the total number of content units by tell and retell, yielded from the content unit analysis (e.g., frame-by-frame descriptions), became the numerator of the equation. These numbers were also totaled to yield a combined tell and retell score. The following formula was then calculated for each participant:

total # words

total # story content units (tell/retell/combined)

After individual lexical efficiency scores were calculated, they were averaged within groups by tell, retell, and combined phases. Lexical efficiency scores for the tell phase were as follows: mean PFCD group score = 70.21 (SD 38.39); mean Comparison group score = 34.08 (SD = 13.4,  $p = .056$ ). Lexical efficiency scores for the retell phase were as follows: mean PFCD group score = 60.17 (SD 25.14); mean Comparison group score = 32.47 (SD = 11.93,  $p = .064$ ). Lexical efficiency scores for the combined tell and retell phases were as follows: mean PFCD group score = 65.19 (SD 31.05); mean Comparison group score = 33.27 (SD = 11.99,  $p = .004$ ). Although the lexical efficiency scores neared significance for both the tell and retell conditions, they did not emerge as statistically significant. However, the combined tell and retell were highly significant ( $p = .004$ ). This finding indicates that the healthy comparison group was significantly more efficient in their overall story telling than the PFCD group.

While Menn et al. (1994) report lexical efficiency norms between 3 – 7 for healthy comparison group participants, it is difficult to relate these norms to our lexical efficiency data due to the differences in operational definitions. More specifically, our definition of content units differed; while Menn et al. (1994) defined content units as correct, informative words, we chose to define content units as the total number of frame-by-frame descriptions the participant's story contained. As such, a reference point for the lexical efficiency scores presented here cannot be provided. However, as illustrated by Menn et al. (1994), lower lexical efficiency scores represent higher communicative efficiency, a trend that is consistent with the data presented here.

An inverse analysis was also run, in which counts of fillers were placed over a denominator of real (intended) words. This analysis was intended to yield measures of dysfluency for each participant. Dysfluencies such as revisions, part-word repetitions, word repetitions, phrase repetitions, and fillers such as *ok*, *uh*, *um*, *you know*, *gosh*, *let me think*, *I mean*, and *ok* were tallied and placed as the numerator. The fillers *um* and *uh* were found to be prevalent throughout all transcripts. Variability within groups was large (e.g., healthy comparison group dysfluency rates ranged from 4.3% to 29.5%, while PFCD rates ranged from 1.6% to 26.5%), and the mean dysfluencies per group were almost identical.

The analysis was re-run, excluding the fillers *um* and *uh*, in order to ascertain whether different patterns of dysfluency would emerge without the most common fillers. Once *um* and *uh* were removed, the mean group dysfluencies differed more-so than in the initial analysis (PFCD group 7.02% dysfluent, healthy comparison group

4.56% dysfluent). Interesting patterns were found within the PFCD group, in which two participants were very fluent (even more-so than their healthy comparison group counterparts), while three participants were very dysfluent. The two fluent PFCD group members (with dysfluency rates of 2.5% and .8%), however, presented with other problems, including syntactic and semantic errors. Potential trade-offs between fluency and narrative structure and content, as well as potential sub-groups within the PFCD patient population, may be fruitful concepts to explore in future work.

### **Descriptive Statistics**

To obtain a general overview of the answers to the remaining study questions, descriptive statistics were computed for content units, thematic units, and story grammar units (See Table 11). On these measures, overall, the stories of the healthy comparison group consistently contained a larger number of content units, thematic units, and story grammar components, and fewer discourse errors, than the PFCD group in both tell and retell phases. Although the ranges of performance overlapped, the healthy comparison group generally performed in higher ranges than the PFCD group. In addition, the standard deviations associated with the performances of both groups were similar in measures of content and thematic units, indicating similar individual variation on these tasks across participants. Greater variation in standard deviations was present on measures of story grammar components.

When examining performance on an individual basis within groups, it is important to acknowledge that even within our small sample, the participants were not normally distributed. One person in each group would be considered an outlier based on their performance (e.g., one healthy comparison group participant

performed similarly to the PFCD group, while one PFCD group participant performed similarly to the healthy comparison group). If this study had contained a larger sample size, it is likely that these participants would have been excluded as outliers.

Table 11. Descriptive Statistics: Content Units (CU), Thematic Units (TU), and Story Grammar (SG)

Descriptive Statistics - Mean (SD) - Range [ ]		
Level - Phase	PFC	Controls
CU - Tell	7.0 (2.55) [3-10]	10.40 (2.608) [7-14]
CU - Retell	7.2 (3.194) [3-12]	9.60 (2.074) [7-12]
CU - Combined	14.2 (5.675) [6-22]	20.0 (4.637) [14-26]
TU - Tell	2.4 (.894) [2-4]	3.40 (1.517) [1-5]
TU - Retell	2.2 (1.095) [1-4]	3.20 (1.304) [1-4]
TU - Combined	4.6 (1.949) [3-8]	6.6 (2.793) [2-9]
SG - Tell	7.20 (1.789) [5-10]	10.40 (4.506) [3-15]
SG - Retell	8.0 (2.345) [6-12]	9.20 (3.899) [3-13]
SG - Combined	15.2 (3.894) [13-22]	19.60 (8.385) [6-28]

### **Between-group Differences: Story Processing Time, Comprehension Question Accuracy, Content Units, Thematic Units, Story Grammar Features, and Discourse Errors**

Because responses violated expectations of homogeneity of variance, between-group differences were examined via a non-parametric measure, the Mann-Whitney Test, for story processing time and accuracy of comprehension questions. Since these measures replicate, to some extent, the more conventional assessments that do not adequately differentiate deficits in PFCD patients, the two groups were not necessarily expected to differ on these measures. As expected, the PFCD group and the healthy comparison group did not demonstrate statistically significant differences on measures of story processing time (in milliseconds) (PFCD Group Mean = 11856.66; SD = 3700.02; Comparison Group Mean = 9612.78; SD = 1092.14;  $p = .117$ ) and accuracy of responses to comprehension questions (PFCD Group Mean =



.77 (e.g. 77% accuracy); SD = .178; Comparison Group Mean = .89; SD = .09;  $p = .280$ ).

Paired t-tests were utilized to assess between-group differences on overall performance on content units, thematic units, and story grammar features. Mean overall content unit scores were as follows: PFCD group = 14.2 (SD 5.675); Comparison group = 20.0 (SD = 4.637). Mean overall thematic unit scores were as follows: PFCD group = 4.6 (SD 1.949); Comparison group = 6.6 (SD = 2.793). Mean overall story grammar scores were as follows: PFCD group = 15.2 (SD 3.894); Comparison group = 19.6 (SD = 8.385). Despite the clear trends in favor of the healthy comparison group, no significant differences emerged.

Paired t-tests were then utilized to analyze between-group differences in discourse errors overall and by phase (tell/retell). The square root of the number of errors in each category was taken prior to analysis in order to normalize the distribution. Because of the low frequency of some error types, some variables were combined, as laid out in the *Transcription of Sample/Coding* section, to create conceptually motivated composite scores. For example, the “total word” category consisted of the total errors made in the nine individual word-level categories, while the total errors at the phrase level consisted of the total errors made in the eight individual phrase-level categories. In addition, the  $p$  value for significance was adjusted to  $p < .025$  to minimize the likelihood of a Type I error, in which a true null hypothesis is incorrectly rejected.

Table 12 consists of a comprehensive listing of  $p$  values for the tell, retell, and combined conditions that were yielded from between-group paired t-tests. The

statistically significant  $p$  values are bolded and underlined. As illustrated in Table 12, multiple  $p$  values reached statistical significance, while others showed a trend toward significance. Two categories (difficulty interpreting gist and faulty coherence inferencing) were theoretically thought to strongly differentiate between the two groups, but did not reach statistical significance. Overall, there were more significant differences between groups when tell and retell data were combined, because of the increase in power due to the increase in number of cases analyzed. In addition, the differences seemed to decrease in significance by phase, as more differences emerged in the tell phase than in the retell phase. Embellishment was the only error category in which significant group differences were found in both the tell and retell phases.

As the “nonaphasic” PFCD population typically is not thought to exhibit many word-finding difficulties, it is not surprising that many of these categories did not emerge as statistically significant. However, word-level errors were noted throughout the transcripts of the PFCD group. Anaphoric cohesion, a word-level error in which a pronoun cannot be referred back to its reference, differed significantly between groups, both in the tell and combined conditions. Further, superordinate substitutions (e.g. “place” for “farm,” “thing” for “plant,” etc.), another word-level error, differed significantly between groups in the tell phase. In addition, PFCD patients exhibited more errors than the healthy comparison group in every word-level category. Overall, although word-level errors were not examined in detail, their presence in the PFCD group transcripts suggests that future research will help to determine whether lexical errors are also characteristic of the discourse of the PFCD population.

Table 12. Discourse Error Analysis

Discourse Errors	Level of Identification	Tell	Retell	Combined
Total Word	Total Word	<b>0.011</b>	0.053	<b>0.001</b>
Total Phrase	Total Phrase	0.168	<b>0.019</b>	<b>0.008</b>
Difficulty Interpreting Gist	Global	1	0.374	0.591
Semantic Paraphasias	Word	0.216	0.15	0.038
Phonemic Paraphasias	Word	0.374	N/A	0.343
Formal Paraphasias	Word	0.374	0.186	0.087
Superordinate Substitution	Word	<b>0.03</b>	0.621	0.081
Unrelated Word	Word	0.621	0.374	0.343
Unretrieved Word	Word	0.178	N/A	0.168
Perseveration	Word	0.374	0.198	0.095
Anaphoric Cohesion	Word	<b>0.015</b>	0.035	<b>0.001</b>
Conjunction Cohesion	Word	N/A	N/A	N/A
Embellishment	Phrase	<b>0.024</b>	<b>0.005</b>	<b>0</b>
Confabulation	Phrase	0.314	<b>0.021</b>	<b>0.016</b>
Perseveration	Phrase	N/A	N/A	N/A
Topic Stray	Phrase	0.374	0.186	0.101
Faulty Predictive Inferencing	Phrase	N/A	N/A	N/A
Faulty Backward Inferencing	Phrase	0.374	N/A	0.343
Faulty Coherence Inferencing	Phrase	0.669	0.854	0.833
Faulty Temporal Sequencing	Phrase	0.198	0.374	0.095

NOTE: Significant p values in bold

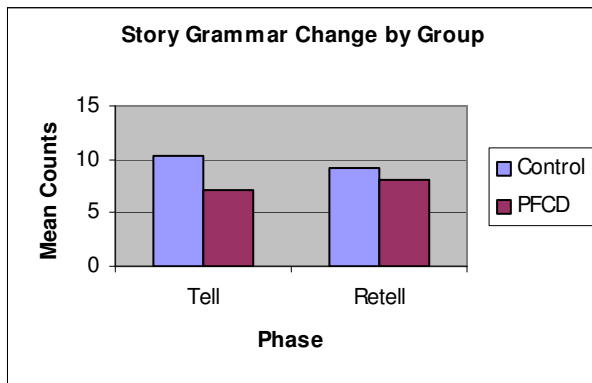
N/A: Not present in either group's transcripts

### Between Group Differences – Tell vs. Retell

Several 2x2-type analysis of variance (ANOVA) were utilized to assess between group differences on the indices of content units, thematic units, and story grammar features, and to examine how and if the tool distinguishes between groups by the tell versus the retell phase. Of the three indices, a significant interaction for group by phase was found for story grammar units ( $p=.0415$ ) (See Figure A). Mean story grammar scores for the tell phase were as follows: PFCD group score = 7.2 (SD 1.79); Comparison group score = 10.4 (SD = 4.5). Mean story grammar scores for the retell phase were as follows: PFCD group score = 8.0 (SD 2.34); Comparison group score = 9.2 (SD = 3.89). Overall, the stories told by the healthy comparison group participants contained more story grammar categories in both the tell and retell

conditions than the stories of PFCD participants. Interestingly, during the retell phase, the PFCD group performed slightly better from a lower baseline score, while the healthy comparison group's performance decreased slightly.

Figure A. Story Grammar Change by Group



No other significant differences emerged; however, trends were in the same direction in the remaining indices of content and thematic units, with the stories of the healthy comparison group containing larger numbers of each in both the tell and retell phases (See Figures B, C, below). Specifically, mean content unit scores for the tell phase were as follows: PFCD group score = 7.0 (SD = 2.55); Comparison group score = 10.4 (SD = 2.61). Mean content unit scores for the retell phase were as follows: PFCD group score = 7.2 (SD = 3.19); Comparison group score = 9.6 (SD = 2.07). Although there was a trend toward both a main effect and an interaction on the content unit analysis, the differences were not statistically significant. Mean thematic unit scores for the tell phase were as follows: PFCD group score = 2.4 (SD = .894); Comparison group score = 3.4 (SD = 1.52). Mean thematic unit scores for the retell phase were as follows: PFCD group score = 2.2 (SD = 1.09); Comparison group score

= 3.2 (SD = 1.3). The number of content units coded in the PFCD group increased slightly from the tell to the retell phase, but the general trend in both groups was a decrease in performance from the tell to the retell phase. It is important to note that it is likely that this measure would not have reached significance if the  $p$  value had been adjusted for Bonferoni.

Figure B. Content Unit Change by Group

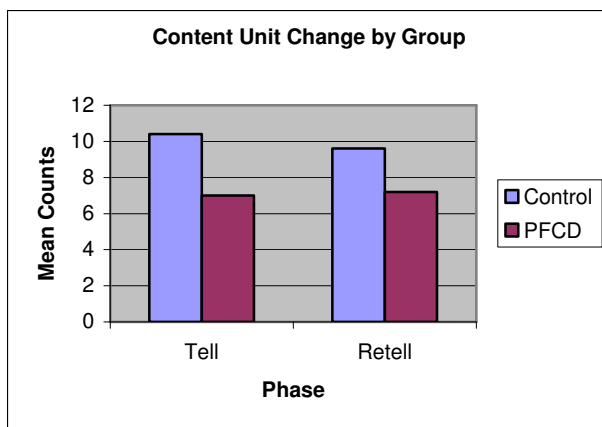
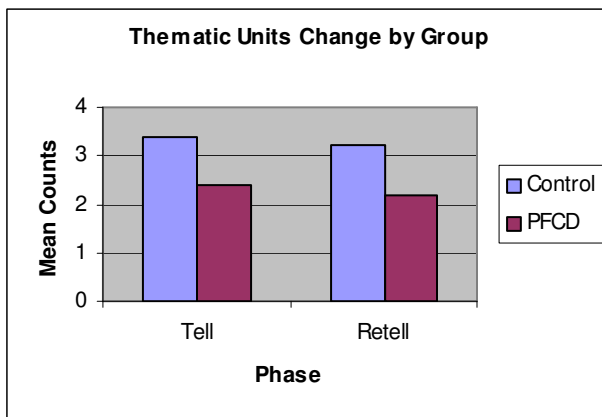


Figure C. Thematic Unit Change by Group



### **Qualitative Analysis –Between-group Differences: Story Grammar**

Story grammar was further explored by category as opposed to overall score.

The story grammar components described earlier included: *setting*, *initiating event*,

*internal response, goal, attempt, outcome, and reactions* (Van Den Broek, 1994). As described under the *Transcription of Sample/Coding* section, in order to code story grammar categories, one story grammar category was identified for each frame of the frame-by-frame description to create a template. These story grammar categories and their associated frame-by-frame descriptions were then compared with each transcript and scored as present or absent. See Table 13 for story grammar counts by category.

Table 13. Story Grammar Counts by Category

	PFC Tell	Control Tell	PFC Retell	Control Retell	Total possible across participants	Total possible per story
Setting	6	6	7	6	10	2
Initiating Event	2	3	3	3	5	1
Goal	2	4	3	4	10	2
Attempt	3	7	5	5	10	2
Outcome	17	25	14	20	35	7
Reaction	6	6	6	6	15	3

Qualitative analysis suggests that the stories of the healthy comparison group tended to contain a larger number of story grammar categories than the PFC group in the attempts, goals, and outcomes categories. In the “attempt” category during the story-tell phase, the healthy comparison group included 40% more components than the PFC group. In the “outcome” category during the story-tell phase, the healthy comparison group included approximately 20% more components than the PFC group. The healthy comparison group also included larger amounts of “goal” statements than the PFC group. This may suggest that as the story became more involved, moving past the initial setting statements and initiating events, the PFC participants experienced more difficulty than the healthy comparison participants.

## **Qualitative Analysis – Summary of Data Trends**

Overall, it should be noted that the stories of the PFCD group consistently contained fewer content units, thematic units, and story grammar features than the stories of the healthy comparison group, although many of these differences did not achieve statistical significance, perhaps due to small sample size. In addition, the PFCD group's stories consistently contained more discourse errors than the stories of the healthy comparison group. For example, in each of the 20 discourse error analysis categories, the comparison group made fewer errors than the PFCD group. In addition, there were multiple discourse error categories that were only coded in the stories of the PFCD group, including: phonemic paraphasias, formal paraphasias, unretrieved words, perseveration at the word level, backward inferencing, embellishment, and topic stray.

Finally, although the literature suggests that the following can be characteristic of the language of PFCD patients, there were several categories in which no errors were made by either group on either the tell or retell (See Table 14), namely: conjunction cohesion, perseveration at the phrase level, and predictive inferencing. There were a few categories motivated by the literature in which no errors were made by either group on the retell only, namely: phonemic paraphasias, unretrieved words, and backward inferencing.

Table 14. Discourse errors not present in either group's sample

Uncoded Discourse Error Category	Phase
Conjunction Cohesion	Tell & Retell
Perseveration (phrase level)	Tell & Retell
Predictive Inference	Tell & Retell
Phonemic Paraphasia	Retell
Unretrieved Word	Retell
Backward Inferencing	Retell

## **DISCUSSION**

### **Overview**

The purpose of this study was to outline the development of discourse codes designed to differentiate between the “nonaphasic” but characteristic discourse production deficits that could present following PFCD and the discourse of a healthy comparison group.

Specifically, this preliminary study was designed to answer the following research questions:

1. Do the coding categories demonstrate sufficient inter-rater reliability?
2. Do the scores derived from the discourse analysis differ significantly between the PFCD group and matched healthy comparison groups on the indices of content units, thematic units, story grammar features, and discourse errors?
3. Do the scores and indices derived from four levels of analysis (content units, thematic units, story grammar components, and discourse error analysis) differ significantly between tell and retell scores by group, with each group performing more poorly on the retell?
4. Does a qualitative error analysis profile differentiate between the two study groups, with the PFCD population demonstrating discourse errors classically



associated with PFCD, and the healthy comparison group exhibiting relatively few such errors?

Overall, findings indicated fair to good inter-rater reliability for the codes, with the variability in agreement dependent in part on the level of analysis (e.g., content units, thematic units, story grammar features, and discourse error analysis). The stories of the healthy comparison group also contained more content units, thematic units, and story grammar features than the PFCD group, despite few statistically significant differences. In addition, the stories of the healthy comparison group contained fewer discourse errors than the stories of the PFCD group. Overall, the participants stories tended to contain fewer components in the retell phase than the tell phase, although no significant differences emerged between groups. Finally, the qualitative error analysis revealed relatively few discourse errors overall by the healthy comparison group, while the PFCD group displayed characteristic discourse errors.

The remainder of the discussion section will be organized in the following manner. First, each of the research questions posed by the study will be explained within the context of the results. Second, discourse models will be examined as applied to the results. Third, the discussion will focus on how the limitations of the study may explain the relative lack of significant results. Fourth, challenges in using narrative analysis as a clinical tool will be discussed. Fifth, future task adjustments will be suggested. Sixth, possible explanations will be provided for the results of the study in light of the extant literature. Finally, implications for future research will be discussed.

## **Do the Codes Demonstrate Sufficient Inter-rater Reliability?**

Narrative discourse is increasingly being used as a clinical tool, in part because it is thought to mimic genuine communication experiences. Thus, reliability is a critical piece of the discussion of any narrative analysis. Results suggest that the discourse codes created for this study (including content units, thematic units, story grammar components, and discourse error analysis) demonstrate fair to good inter-rater reliability when coding was completed by multiple independent volunteer judges. Both a Kappa statistic for categorical data and an Interclass Correlation Coefficient (ICC) were performed to examine this research question. The Kappa statistic findings indicated that agreement for content units, thematic units, and story grammar units was within the fair to good range, with the exception of poor thematic unit agreement between two judges. ICC findings for the discourse error analysis suggested poor to excellent agreement, with the poor agreement emerging on the judgment of adequate/inadequate story gist. Although the final data for content units, thematic units, and story grammar components was determined by the majority, it should be noted that the two judges trained to make reliability judgments were less sensitive overall than the primary judge (primary author IE). In other words, the two trained judges detected fewer content units, thematic units, story grammar components, and discourse errors than the primary judge (primary author IE). Very similar inter-rater agreement was found for story grammar components and content unit analysis. This is of interest because the story grammar analysis theoretically tapped story structure along with story content, while the content units only tapped story content.

## **Reliability – Suggestions Relative to the Training Process**

Despite the presence of fair to good agreement overall, better agreement was expected for the content unit analysis due to the relatively specific nature of the frame-by-frame descriptions. It appears that identifying the presence or absence of a frame in a story-tell transcript is more complex than initially anticipated. In the reliability training manual provided to the judges, one concept was stressed repeatedly: “Do not expect the stories of the participants to match the frame descriptions verbatim – just try to determine whether or not the information is adequately mentioned.” Although the content unit analysis seemed relatively objective, the subjective nature of determining “adequate mention” may need to be clarified in the future, in order to improve reliability ratings. One strategy may be to set some numeric quantifier, such as needing to find greater than 50% of the frame description in the transcript in order to code it as present.

Other modifications to improve the training procedure may include coding several sample transcripts across all four levels of analysis with the judges-in-training. In addition, it is possible that some of the operational definitions should be changed, or more examples added, in order to increase the sensitivity of the training/coding process. For example, inter-rater reliability of the discourse error analysis was within the good to excellent range for word and phrase-level errors, but was quite poor for the “difficulty interpreting gist” measure. One suggestion for future revision of this tool would be to provide multiple examples of sentences or ideas that would convey a “partial or narrow ability or inability to capture theme or gist” (Frattali & Grafman, in press), as currently defined in the training manual.

## **Reliability – Suggestions Relative to the Task**

Similar issues emerged when examining agreement for thematic units, in which the subjective nature of the coding process was even more pronounced. Here, judges were asked to determine whether five themes were present or absent in the story transcripts. Not only was inter-rater reliability influenced by the fewer number of judgments to be made (and the small  $N$ ), but also by the subjective nature of examining an entire transcript for five relatively abstract ideas. One way in which to possibly circumvent this issue would be to ask the participants, subsequent to story presentation, to describe the themes of the story. That way, judges could determine whether or not those themes matched the predetermined ones from a more restricted context.

A revision to the task could also include asking the participants, subsequent to their story-tell, to express the gist or single overarching theme of the story, which would provide a somewhat more objective means for judgment, possibly enhancing the reliability. It should be noted that the more extensive training that is needed impacts ease of clinical use in addition to reliability. For childhood language assessments, this issue has been recognized, and clinical tools that employ narrative contain extensive training manuals (e.g., Strong, 1998).

## **Between-group Differences: Content Units, Thematic Units, Story Grammar Features, and Discourse Feature Errors**

The results of this study suggest that between-group differences exist across all four levels of analysis, despite relatively few statistically significant findings. Paired  $t$ -tests were utilized to analyze between-group differences in content units, thematic units, and story grammar features. No statistically significant differences

emerged in any of these three levels of analysis, although the stories of the healthy comparison group consistently contained more substance than the PFCD group. Given a larger sample size, it is likely that what was only a trend would reach statistical significance. Also, task instructions may have been somewhat misleading. Participants were asked to “Tell the story now, from beginning to end. Be as specific as possible.” Asking participants to be as specific as possible may have led them to focus on details as opposed to overall story content and themes. Adjusting the instructions may lead to more distinct differences between groups.

Paired  $t$ -tests were utilized to examine between-group differences in discourse errors. Statistically significant ( $p < .025$ ) between-group differences were found for multiple discourse error categories, including total word, total phrase, embellishment, confabulation, superordinate substitution, and anaphoric cohesion. Two additional categories that were thought theoretically to strongly differentiate between groups did not reach statistical significance (difficulty interpreting gist and faulty coherence inferencing). This lack of significance may be due to unclear operational definitions, and, in the case of difficulty interpreting gist, the complexity of making a single, subjective judgment as to the theme of the overall story. (Suggestions regarding how to modify and better tap story gist were provided under the *Inter-rater Reliability* section of the discussion.)

Several discourse errors classically associated with PFCD patients were included in the discourse error analysis and yet not identified in any transcripts. These included errors in conjunction cohesion, perseveration at the phrase level, and faulty predictive inferencing. The operational definition of conjunction cohesion is:

“an error in the type of conjunction tie used to link utterances.” This type of error, although discussed in the literature as a possible characteristic of the PFCD population, is not considered a core characteristic. Future research will determine whether it should remain in the error analysis.

Perseveration was listed twice in the error analysis: once at the word level, and once at the phrase level. It was only coded by the judges at the word level, which may indicate difficulty differentiating between the two operational definitions. For future use, either the operational definitions should be clarified, or perseveration should only be listed once, to decrease confusion.

Finally, faulty predictive inferencing was not noted in any transcript, which was most likely due to the nature of the story tell/retell task. Faulty predictive inferencing was defined as “An error in forward inferencing that reflects the readers’ inability to anticipate and/or predict a future event.” The likelihood of making such an error when telling a story after having viewed the entire story is slim, and as such, should be removed from the error analysis of this specific task unless the overall structure of the task undergoes revision.

### **Between-group Differences – Tell vs. Retell**

A 2 x 2-type ANOVA was utilized to determine whether the tool distinguishes between groups by the tell versus the retell phase. Results suggest that a statistically significant interaction of group by phase emerged on story grammar components only. However, certain trends were noted, namely, that overall, performance decreased from the tell to the retell phase in both groups. When analyzing story grammar in more detail, it was found that the interaction was due to the slight

improvement in the PFCD group's performance from tell to retell in story grammar components. Because the PFCD group started off with fewer story grammar components than the healthy comparison group, they theoretically had less to lose in the retell phase. It is also possible that repetition is potentially a valuable learning strategy for these patients.

It is essential now to examine the ramifications of utilizing a story tell/retell construct with this population. It has been noted in the literature that stories told from memory tend to be longer in both words and utterances and contain more errors than stories told while looking at a picture (Morris-Friehe & Sanger, 1992). In addition, Morris-Friehe and Sanger (1992) suggest that stories told from memory may require more complex manipulation and development of story structure than other story elicitation tasks. Although in the case of their study, the story told from memory was fictional and not related to pictorial stimuli, these same notions of increased complexity may apply to the story tell/retell construct utilized in this study. The story tell/retell construct appeared to be the logical choice for our task, because the story was relatively complex in and of itself, and it was thought that the participants might experience difficulty truly understanding the story until the entire story was viewed. Thus on the one hand, the complexity of the story decreased with a full viewing prior to story-telling, but the demands of the actual story-telling increased with the task set-up. Morris-Friehe and Sanger (1992) also suggest that stories yielded from picture tasks may represent more "typical" and less complex levels of language. Because this study intended in some way replicate the type of language seen in day-to-day living,

it is possible that a task involving telling the story while viewing the pictorial stimuli may have yielded more functional results.

In relating the story tell/retell construct to the results of this study, the PFCD group often presented with fairly little in the way of content, themes, and story grammar components in their initial story-tells. As such, they had relatively little to lose in their retells. In contrast, the healthy comparison group presented with more substantial content, themes, and story grammar components in their initial story-tells. They had more to lose when it came to retelling their stories, and as such, moved closer to the relatively stable PFCD group scores in the retell. This pattern applied to the discourse error analysis as well, in which more categories reached statistical significance between groups in the tell phase than in the retell phase, suggesting that the number of errors made in the retell were more similar between groups.

Another factor that may have played into the more similar performances in the two groups in the retell phase is memory. It is possible that the set-up of the task, in which the story was told immediately after the viewing and then again after a 30-minute time lapse in which the participants were involved in other protocol activities, may have affected the retell of both groups. In order to assess the risk of memory as a confounding factor, the PFCD group was given the *Wechsler Memory Scale – 3* (Wechsler, 1997), on which only one participant performed in a clinically significant range. This participant's performance may have been impacted by memory weaknesses, but all other PFCD group members performed adequately on the *Wechsler Memory Scale – 3* (Wechsler, 1997). As such, it is likely that most participants were affected similarly by the 30-minute interval in terms of memory



degradation. However, in the future, it may be helpful to assess the memory functions of the healthy comparison group as well.

## **Qualitative Analysis – Between-group Differences**

The results of this study suggest that qualitative analysis does distinguish between groups and may be exploited in future diagnostic instruments. The healthy comparison group made very few errors in the discourse error analysis, and the stories of the healthy comparison group consistently contained more content units, thematic units, and story grammar components than those of the PFCD group. Additionally, the stories of the PFCD group were found to contain multiple discourse error categories that the healthy comparison group did not contain, including: phonemic and formal paraphasias, unretrieved words, perseveration at the phrase level, backward inferencing, embellishment, and topic stray.

Qualitative analysis also allows for an examination of the clinical significance of the different profiles that presented. To illustrate the qualitative differences between the PFCD group transcripts and those of the healthy comparison group, excerpts were taken from two story-tell transcripts (See Table 15).

Table 15. Excerpts from participant transcripts

Excerpt from PFCD group member transcript:

Old Macdonald had a *had a had a* farm. It's *um* he and his wife. And he *um* he *um*.... Gosh, let's start with her *let's start with her um* efforts. He *um* she has *um* a little flower that blooms. See I'm doing nothing from April. And then all of a sudden it blooms. And then in May it just blossoms into a new *a new uh* sprout of *um*.... You know *you know* what *is is*.... A small *um* herb or something growing in the kitchen. *Um* he *um* has *um* he-'s a superintendent of a building. That is *um that is um* he-'s superintendent of a building. And it's renovated *uh um*.... Decided to *um decided to um* grow crops in every different room that he has in the *in the* building. (14 utterances; 135 words total; 117 without fillers as calculated under lexical efficiency section).

Excerpt from healthy comparison group member transcript:

And he and the husband also started to plant all these vegetables out where he chopped down the tree. And he also had cows in there, too, in this apartment building. And a lot of people who lived in this apartment building were getting upset because it kept growing and growing. And he had carrots that were coming out of the ceiling in a different apartment... And things just grew, and the plants ended up just taking over the whole apartment building. A lot of it looks like that a lot of the residents of the apartment building were upset about it and ended up moving out. (6 utterances; 107 words total; 107 words without fillers as calculated under lexical efficiency section).

Although multitudes of clinically significant differences may not have emerged in data analysis, it becomes very clear when examining these excerpts that there are distinct differences between the story-tells. The PFCD patient repeats, perseverates on extraneous information, revises, strays from the topic, has difficulties retrieving certain words, uses other words incorrectly, has difficulty sequencing, exhibits multiple sentence fragments, and takes many utterances to express little relevant story content (exhibits poor lexical efficiency). In contrast, although the healthy comparison group member does make some errors in anaphoric reference and often revises thoughts at sentence initiation, this participant expresses significantly more story content with less than half the number of utterances and fewer words (good lexical efficiency). This brief excerpt analysis also highlights the relevance of lexical efficiency in examining and comparing discourse-level transcripts of the PFCD group and healthy comparison participants. In addition, it showcases the overall differences in fluency that can present between the two groups, although dysfluency levels were not consistent within groups. Through this illustration, it becomes clear that a story elicitation task can be an effective clinical tool that evokes many of the characteristic discourse-level errors of PFCD patients.

Examining content unit analysis, thematic unit analysis, and story grammar feature analysis from a clinical perspective also yields interesting insights. The five PFCD group members consistently expressed less story content, fewer thematic units, and fewer story grammar components than the healthy comparison group. However, the results of all three analyses for both groups ended up being in the middle of the distribution. More specifically, across these three analyses, the PFCD group scored in the mid-40<sup>th</sup> percent range, while the healthy comparison group scored in the mid 60's. Because means overlapped, these three analyses in and of themselves are not as clinically useful as the discourse error analysis. However, when completed as a whole, all four levels of analysis could contribute to a clinical profile. Further, the task will likely be of greater clinical utility subsequent to the adjustments suggested throughout this discussion. For example, a faster and easier way to tap a patient's grasp of the themes and gist of the story would be to simply ask them after they have finished their story-tell and retell, and then to compare these answers to a template. When this is completed, along with a story grammar analysis and a discourse error analysis, a clinically relevant profile would emerge, and would provide the clinician with information that could provide the basis for treatment on the discourse level.

Finally, a more in-depth qualitative analysis of story grammar was completed. Healthy comparison group participants provided more goal, attempt, and outcome statements than the PFCD group participants. This pattern of performance differs somewhat from literature from various patient populations thought to present with fairly similar profiles. For example, the literature suggests that the student-aged learning disabled population makes significantly fewer statements than normally

achieving students in the response, attempts, and planning categories (Roth & Spekman, 1986). Less research has been conducted examining the TBI or RHBD populations specifically utilizing the story grammar schema (Snow et al., 1999). However, some literature suggests that even severely brain injured patients may have preserved story grammar knowledge subsequent to their injury (Snow et al., 1999). More research exists examining episode use and structure in the TBI population, which suggests that TBI and non-brain injured (NBI) participants do not differ in terms of total number of episodes produced; however, NBI participants produce fewer T-Units that don't contribute to episodic structure (Coelho, 2002). The application of story grammar schemas to the TBI, RHBD, and PFCD populations remains sparse, and future research along these lines is warranted.

### **Discourse Models as Applied to Results**

The discourse analysis that was developed for this study was cast within features of Grafman's (2002) framework of PFC function, which describes features such as temporal sequencing of events, cohesion, coherence, and gist as localized representationally within the PFC. The results of this study appear to support the notion that features such as temporal sequencing, cohesion, and coherence are impaired with damage to the PFC, as the PFCD group experienced qualitatively more difficulty with them than the healthy comparison group, although the differences were not statistically significant. As previously mentioned, gist, as operationalized and measured in this study, did not differentiate between groups. However, with the modifications suggested for the gist measure, it is likely that differences will emerge in the future. In addition, given the trends in performance, it is likely that many of the

differences mentioned above would have been statistically significant with a larger sample size.

## **Study Limitations**

There are several limitations to this study that could explain the relatively low level of statistical significance emerging from the data analysis. First, the sample size of 5 in each group was quite small, even for the basic statistical analyses conducted. Clear trends in group differences emerged throughout that suggest that a larger sample size would have yielded more statistically significant results. In addition, when tell and retell data were combined, more statistically significant results surfaced. Therefore, it seems probable that a larger  $N$  would elicit more statistically significant between-group differences on all measures.

Another study limitation is related to lesion location and volume. All five PFCD group participants were fairly mild in their lesion size/volume. As such, their presentations were also fairly mild on the PFCD spectrum. It is possible that this contributed to the relatively little statistical significance when looking at between-group differences. Further, it is possible that milder-lesioned patients do not show pronounced deficits in the discourse areas tapped by this study. In addition, the PFCD group's lesion locations, although confined to the PFC, ranged from the right, left, and bifrontal regions. In the future, it would be preferable if all PFCD group members had a lesion confined to one hemisphere (e.g. either right or left).

The next study limitation is related to the task and memory deficits. As discussed earlier, one PFCD group member presented with low, clinically significant scores on the *Wechsler Memory Scale – 3* (Wechsler, 1997). This may have skewed

their results, and therefore the results of the study at large, as the task required the use of short-term memory.

Finally, *Old MacDonald had an Apartment House* (Barrett & Barrett, 1969) has been utilized very little in the narrative literature. Other stories with varying length, complexity, content, style, or themes may have tapped the discourse deficits of PFCF patients more effectively. However, the primary narrative elicitation task used in the child language literature is *Frog, Where Are You?* (Mayer, 1969), which may not be sufficiently complex for use with the PFCF population.

### **Challenges in Using Narrative Analysis as a Clinical Tool**

Various challenges have been noted in the literature regarding the use of narrative analysis as a clinical tool. First, it is generally time-consuming and as such can be difficult for busy clinicians to fit into their assessment batteries. In addition, because narrative analysis, especially at the level of discourse, is often intended to examine the more subtle aspects of language, it can be difficult to create and utilize appropriate operational definitions (Mentis & Prutting, 1991). While the searched-for subtleties are often intuitively understandable, they can be difficult to describe objectively (Brinton & Fujiki, 1989).

This subjectivity then influences reliability, which can be very difficult to achieve. Judgments, especially those of “appropriateness,” are subjective in nature and can easily be impacted by the judge’s outlook and approach (Adams, 2002). The difficulty achieving satisfactory reliability is further influenced by the sheer quantity of narrative features that can be tapped, which can be quite high (as it was in this study). It may also be influenced by the possible variability of features studied, even

within patient population groups. More structured narrative tasks are often considered advantageous to spontaneous story generation tasks because they are easier to administer and analyze (Adams, 2002). However, even when utilizing a more structured narrative task, as in this study, the variability, even within groups, was startling. In addition, as in any narrative task, the analysis is constrained by what is yielded. It is also difficult to discern whether narrative tasks mimic genuine communication experiences, although they are certainly more efficient than naturalistic observation. However, in either approach, the data may not accurately reflect the abilities and knowledge of the participants (Adams, 2002).

### **Suggested Task Adjustments**

For future use, it may be helpful to modify the tell/retell task in order to increase its effectiveness in capturing the deficits of PFCD patients. Some task and training modifications have been suggested throughout the discussion. The following consist of some additional suggestions. First, participants should tell the story as they are viewing it pictorially, and then retell it once they have finished viewing the entire story, to decrease the likelihood of memory-based influences. In order to streamline the task and to reduce redundancy, story grammar analysis and content unit analysis should be further differentiated. For example, while content units would remain as a strict measure of story content, the story grammar analysis could analyze both by category and by episode integrity (completeness).

Next, although lexical efficiency should continue to be calculated, the types of fillers excluded in this measure should be expanded (e.g., also exclude “you know,” “gosh,” etc.), to enhance sensitivity. In addition, lexical efficiency as a measure may

actually tap various problems that yield similar outcomes. For example, self-corrections, word fragments, and circumlocutions may all be the result of word-finding problems, whereas digressions and irrelevant statements may be caused by other factors. As such, lexical efficiency may be further differentiated into several categories to discern whether certain types of differences would be more common in the PFCD group than in the healthy comparison group. This differentiation may include a dysfluency category, as dysfluencies can decrease coherence due to their interference in message transmission (Bliss, McCabe, & Miranda, 1998). Also, because narrative is open-ended, and multiple words can be used to describe the same narrative component, the word choices of each group could be analyzed. For example, the PFCD group might opt towards simpler (shorter/higher-frequency) words (e.g., animal), and avoid more precise terminology (e.g., cow). While this is similar to the “superordinate substitution” category of the discourse analysis, it may be captured more effectively in other ways.

Another suggestion involves narrowing down the twenty variables included in the discourse error analysis. Several key variables emerged in this study, including embellishment, confabulation, and anaphoric cohesion. As this tool is utilized and improved upon in the future, an attempt should be made to include only those variables that contribute to a clinically relevant profile, and that are found to distinguish between groups.

A final suggestion relates to the assessment battery administered to the participants. In the future, both a confrontation and discourse naming task should be



administered to the participants, in order to explore lexical issues such as word retrieval vs. discourse-level organization.

## **Study Results in the Context of Current Literature**

The results of this study are consistent with current literature on several levels. First, PFCD patients have been described as having largely intact microlinguistic abilities with impaired suprasentential, discourse-level abilities (Ferstl et al., 1999). This appears to be the case with the PFCD group studied here, although they did display more word-level errors in the discourse error analysis than their healthy comparison group counterparts. Generally, however, the microlinguistic impairments did not consist of sheer word-finding or paraphasic errors; rather, they consisted of superordinate substitutions and faulty anaphoric references. Faulty anaphoric reference has been classically associated with PFCD (Frattali & Grafman, in press), so the significant group differences on this measure were not surprising. However, the relative prevalence of word-level errors should be further examined in future research.

As there have been relatively few studies describing the effects of prefrontal cortex damage on discourse (Alexander, 2002), this area merits further exploration. However, the literature suggests that characteristics associated with PFCD patients' discourse production, specifically in the context of story-telling, include: difficulty recalling narrative components of a story, processing inference and appreciating the story's thematic aspects or gist (Frattali & Grafman, in press; Zalla et al., 2002); confabulation; embellishment; topic stray; faulty anaphoric reference and links; faulty temporal sequencing of events and cause/effect relations (Craig & Frattali, 2000;

Frattali & Grafman, in press; Ferstl et. al, 1999); loss of moralistic meaning (Zalla et al., 2002); misinterpreting abstract or implicit information; and producing story tell/retells that either contain intrusive detail or lack detail (Frattali & Grafman, in press). These types of higher-level functions have been associated with the PFC in recent research, and are thought to relate to the PFC's ability to integrate input from various sources and implement abstract behaviors. Overall, quantitative and qualitative data from this study support the presence of these types of characteristics in the discourse of PFCD patients more-so than in their healthy comparison group counterparts.

### **Implications for Future Research**

The results of this preliminary study suggest a number of different avenues for further research. First, the set of tools created for this study (e.g. content unit, thematic unit, story grammar component, and discourse error analyses) should continue to be modified in order to enhance their specificity and sensitivity to the discourse deficits that may be present following prefrontal cortex damage. Second, the tools should be utilized with a larger sample in order to further explore the trends witnessed in this study. It would also be helpful to utilize a more homogeneous PFCD group sample, for example, only accepting PFCD group members with a left PFC lesion. In addition, all study participants should receive scores within the average range on a memory task, such as the *Wechsler Memory Scale – 3* (Wechsler, 1997), prior to admittance to the study. This reduces the risk of memory impacting task performance. Once the set of tools and task are modified and administered to larger sample sizes, measures of inter-rater reliability, internal consistency, and

external validity should be completed in order to begin the process of validating the tool for use by clinicians at large.

Eventually, it is hoped that these tools will be used to assess the discourse not only of patients with PFCD, but other patient populations with discourse-level deficits as well, including but not limited to: schizophrenia, epilepsy, learning disabilities, TBI, and RHBD. The tools created here are eventually intended to assist in the diagnosis of these patients, provide a basis for therapeutic intervention, and track changes in discourse over time.

## **CONCLUSION**

The purpose of developing the series of four analyses was to pilot a set of quantitative and qualitative measures that would capture the essential and fundamental characteristics associated with the discourse of PFCD patients, as these characteristics have been defined clinically. On the basis of the data presented above, the results of this study support the original aims. The findings of this study indicated fair to good inter-rater reliability for our codes, stronger performances by the healthy comparison group across all four levels of analysis, poorer performance overall on the retell phase than the tell phase, and relatively few qualitative discourse errors associated with the healthy comparison group, with classic discourse errors associated with the PFCD group. Future research will need to enhance the strength of the analyses run and continue to explore ways in which to increase the specificity and sensitivity of the tools created.

## **APPENDIX A – Discourse Analysis Systems**

### **Automated Systems**

#### **CLAN**

##### **Name**

CLAN (Computerized Language Analysis); designed for CHILDES (Child Language Data Exchange System) (MacWhinney, 2000)

##### **Properties**

- Designed to analyze data that are transcribed in “CHILDES” format.
- Graphic interface.
- Completes analyses via a series of commands that search for strings and compute various indices.
- In general, contains various automatic analyses of transcribed data, including frequency counts, word searches, co-occurrence analyses, etc.
- Can create original codes.

##### **Uses**

- Used primarily with child language data, but has been used with adults.

##### **Strengths/Weaknesses**

###### Strengths:

- Codes can be added/manipulated.
- Flexible.

- Performs automatic analyses.
- Via “transcriber mode,” can link digitized data to the transcript itself.
- Can import files from SALT into CLAN with the code “SALTIN.”

Weaknesses:

- Transcription must be in “CHILDES” format.
- Primarily contains purely linguistic codes.
- Discourse codes already in system are interactional and not applicable to monologue.

**Codes/features applicable to PFCD population**

- Codes that apply to discourse deficits of those with PFCD can be created. For example, an error analysis system can be created, in which features in various categories are noted and counted.
- The following CLAN programs will likely be helpful in analyzing the basic features of narrative transcripts:
  - The “FREQ” program counts frequencies by calculating the number of times a word occurs in a file or set of files. It produces a set of all words in the specified file and their frequencies, along with a type-token ratio (TTR).
  - The “MLU” program calculates the MLU of an entire file or a specified subset of a file. MLUw (words) can also be calculated, which doesn’t require breaking words down into morphemes.
  - The “RELY” program checks reliability by spotting matches and mismatches of two (or more) coders coding one file or a group of files.

## **SALT**

### **Name**

Systematic Analysis of Language Transcripts (SALT) (Miller & Chapman, 1985)

### **Properties**

- Performs numerous automatic analyses of transcribed data.

### **Uses**

- Used primarily with child language data.

### **Strengths/Weaknesses**

#### Weaknesses:

- “Closed-system,” meaning that the codes cannot be easily manipulated.

Codes from other systems (i.e., CLAN) cannot be imported into SALT.

## **Manual Systems**

### **Clinical Discourse Analysis**

Damico, 1985

### **Properties**

- Consists of four major parameters: quantity, quality, relation, and manner.
- Within each parameter are certain “qualities.”
- Quantity: insufficient information bits; non-specific vocabulary; informational redundancy; need for repetition.
- Quality: message inaccuracy.
- Relation: poor topic maintenance; inappropriate speech style.

- Manner: revision behavior; linguistic non-fluency; inability to structure discourse; inappropriate intonation contours.
- Summary: total utterances; total discourse errors; total utterances with errors; percentage utterances with errors.

### **Uses**

- Designed to identify language impairments in older school-age children.
- Is meant to be used as a descriptive tool to analyze conversation.

### **Strengths/Weaknesses**

#### Strengths:

- Has been applied to analysis of procedural discourse of patients with TBI (Snow, Douglas, & Ponsford, 1999).
- Can be applied to various other populations

#### Weaknesses:

- No published reliability or validity data
- Coding descriptions are vague (the parameters and qualities are minimally described, with few examples).

### **Codes/features applicable to PFCD population**

- All syllables are classified as: essential, optional, mazes, or low content (repeated or irrelevant information).
- Features relevant to PFCD discourse are as follows:
  - Quantity: insufficient information bits; non-specific vocabulary; informational redundancy.
  - Quality: message inaccuracy.

- Relation: poor topic maintenance.
- Manner: revision behavior; inability to structure discourse.
- Summary: total utterances; total discourse errors; total utterances with errors; percentage utterances with errors.

## **Cohesion Analysis**

Mentis & Prutting, 1987; Halliday & Hasan, 1976

### **Properties**

- Cohesion is analyzed by tallying total sentences, total cohesive ties, and mean ties per sentence.
- Cohesive ties are placed into one of 6 cohesion categories: lexical, reference, conjunction, ellipsis, substitution, and incomplete.

### **Uses**

- Designed to quantitatively and qualitatively compare cohesion abilities of closed head injured and normal adults, and also to compare cohesion abilities in narratives as opposed to conversational discourse.

### **Strengths/Weaknesses**

#### Strengths:

- Examined both conversational and narrative discourse.
- Used with patients with closed head injury.
- Contains adequate interrater reliability for cohesion analysis.

### **Codes/features applicable to PFCD population**

- Reference, ellipsis, conjunction, lexical, and incomplete cohesion may be applicable to the PFCD population.



- Lexical cohesion signals continuity of meaning in text. Therefore, this measure indirectly provides a systematic measure of topic maintenance.

## **Intonation unit analysis of conversational discourse**

Wozniak, Coelho, Duffy, & Liles, 1999

### **Properties**

- Employs a modified version of the intonation unit analysis developed by Mentis & Prutting (1991).
- Conversational samples are segmented into intonation units, and then each intonation unit is placed into an ideational intonation category.
- Ideational intonation categories include: contains new information, no new information, incomplete, or tangential.

### **Uses**

- Designed to comprehensively investigate topic maintenance in unspecified conversational contexts.
- More specifically, the authors wondered whether this analysis would differentiate high functioning CHI patients from controls, and whether a pattern of conversational performance would emerge.

### **Strengths/Weaknesses**

#### Strengths:

- Used for patients with closed head injury.

#### Weaknesses:

- Did not find significant differences between subject groups (possibly due to method of conversation elicitation).

- Only utilized an unspecified conversational context.

### **Codes/features applicable to PFCD population**

- This modified analysis did not successfully differentiate between patients with CHI and controls.

## **Linguistic Communication Measure**

Menn, Ramsberger, & Helm-Estabrooks, 1994

### **Properties**

- Used with the “Cookie Theft Picture” from the Boston Diagnostic Aphasia Examination (BDAE).
- Quantifies the amount of information that a patient can convey verbally, the proportion of informative to non-informative words produced (Index of Lexical Efficiency; ILE), and the grammatical acceptability of speech (Index of Grammatical Support; IGS).
- The ILE increases with jargon, word-finding problems, and verbal paraphasias.
- The IGS is lowered by omissions and morphological errors.

### **Uses**

- Developed to track ability of fluent and non-fluent aphasics to produce oral narratives.
- Designed as a clinical tool to measure “verbal communicative effectiveness” (Menn et al., 1994, p. 345).
- Intended use: tracking changes in individual patients across time.

## **Strengths/Weaknesses**

### Strengths:

- Easy to learn/use.
- Easy to apply to other types of narratives; however, the reliability/validity measures would no longer apply.

### Weaknesses:

- Relative ambiguity in determining what is considered “informative” for ILE.
- Provides little psychometric validity/reliability information, and what little is provided does not suggest that this measure is particularly sensitive.

## **Codes/features applicable to PFCD population**

- Total number of words are broken down into content units. The correct words in the content units are then counted and divided into the correct number of grammatical morphemes.
- The ILE is calculated by dividing the total # of words into the total # of content units. If this number equals 1, every word was informative.
- The IGS is calculated by dividing the total # of correct words in content units plus the number of correct endings into the # of content units. This may or may not be informative for the PFCD population.

## **Quantitative Analysis of Aphasic Sentence Production**

Rochon, Saffran, Berndt, & Schwartz, 2000; Saffran, Berndt, & Schwartz, 1980

### **Properties**

- Focuses on syntactic construction.

- Examines frequency of occurrence of various features, including propositional utterances, elaboration, grammatical morphemes, etc.
- Analyzes lexical, morphologic, and structural information (see Rochon et al., 2000, for details).
- Narrative sample is broken down into narrative “core” of words, and then divided into utterances, which are either designated as a sentence or a subsentence structure.
- Words are placed into various lexical categories, including open class words, nouns, nouns requiring determiners, pronouns, etc.
- Morphological content and structural complexity can also be calculated.
- The measures that make the clearest distinction between the clinical population and the controls include speech rate, determiner index, proportion of words in sentences, the well-formedness measure, the sentence elaboration score, and the median length of utterance.

### **Uses**

- Designed to “capture and describe the speech production patterns” of agrammatic patients (Rochon et al., 2000, p. 193).
- Performed on narrative speech.

### **Strengths/Weaknesses**

#### Strengths:

- Overall, the QPA seems to be a reliable means of analyzing the sentence production of fluent and non-fluent aphasic patients.
- Test-retest reliability relatively high across most scores.

- High interrater reliability for normals and controls, both for assigning utterances and scoring.
- Provides objective means of comparing across patients and tracking change in a single individual across time.

Weaknesses:

- Utilizing the QPA can be difficult, especially for those with lesser degrees of syntactic sophistication.

**Codes/features applicable to PFCD population**

- Analysis of syntactic construction is not applicable to the PFCD population.

**Systemic Functional Linguistics**

Halliday, 1994; Togher, 2001

**Properties**

- Includes analysis of politeness markers, exchange structure analysis, generic structure potential (GSP) analysis, and cohesion analysis.
- Links language and context via three concepts: field, or the activity; tenor, or the participants; and mode, or the role of language in a particular social situation.

**Uses**

- Has been applied to analysis of aphasia, Alzheimer's, and TBI.
- Contains analyses intended for use with interactional discourse.

**Strengths/Weaknesses**

Strengths:

- Contains cohesion analysis.

### **Codes/features applicable to PFCD population**

- GSP analysis (the genre or overall structure of an interaction) is potentially relevant to the PFCD population and task. The specific structural components of a genre are determined by the field (e.g., activity) and tenor (e.g., participants) of the interaction. Narratives are also considered genres.
- Cohesion is also considered an SFL analysis.

### **Topic Analysis**

Mentis & Prutting, 1991

#### **Properties**

- Topic and subtopic maintenance are analyzed for conversation and monologue. Intonation categories are assigned and then broken down into textual intonation units, interpersonal intonation units, and ideational intonation units.
- Ideational intonation units are further analyzed as containing new information, no new information, side sequence units, and problematic ideational units.

#### **Uses**

- Designed to be sensitive to patterns and problems in topic management, which is necessary to establish coherent discourse.
- Contains different topic analyses for conversation vs. monologue.

## **Strengths/Weaknesses**

### Strengths:

- Evidence suggests that this multidimensional topic analysis is reliable in both evaluating topic management and identifying differences between populations.
- Has been used with patients with CHI (Mentis & Prutting, 1991).

### Weaknesses:

- Features were described in relatively vague terms and are therefore difficult to operationally define.

## **Codes/features applicable to PFCD population**

- Can analyze the comprehensiveness of a monologue by examining the number of separate issues (new, unrelated, or reintroduced) pertaining to the monologue topic that is introduced by the speaker.
- The monologue conditions utilized in this article included talking about concrete (e.g., “describe a visit to the dentist”) and abstract (e.g., “tell me about truth”) topics; however, the topic analysis can likely be applied to various tasks.

## **Appendix References**

Damico, J.S. (1985). Clinical discourse analysis: A functional approach to language assessment. In C. S. Simon (Ed.), *Communication skills and classroom success*. London: Taylor & Francis.

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

- Halliday, M.A.K. (1994). *An introduction to functional grammar* (2<sup>nd</sup> ed.). London: Edward Arnold.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Menn, L., Ramsberger, G., & Helm-Estabrooks, N. (1994). A linguistic communication measure for aphasic narratives. *Aphasiology*, 8, 343-359.
- Mentis, M. & Prutting, C.A. (1987). Cohesion in the discourse of normal and head-injured adults. *Journal of Speech and Hearing Research*, 30, 88-98.
- Mentis, M. & Prutting, C.A. (1991). Analysis of topic as illustrated in a head-injured and a normal adult. *Journal of Speech and Hearing Research*, 34, 583-595.
- Miller, J. & Chapman, R. (1985). *Systematic analysis of language transcripts*. Madison, WI: University of Wisconsin.
- Rochon, E., Saffran, E.M., Berndt, R.S., & Schwartz, M.F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language*, 72, 193-218.
- Snow, P.C., Douglas, J.M., & Ponsford, J.L. (1999). Narrative discourse following severe traumatic brain injury: A longitudinal follow-up. *Aphasiology*, 13, 529-551.
- Togher, L. (2001). Discourse sampling in the 21<sup>st</sup> century. *Journal of Communication Disorders*, 34, 131-150.
- Wozniak, R.J., Coelho, C.A., Duffy, R.J., & Liles, B.Z. (1999). Intonation unit analysis of conversational discourse in closed head injury. *Brain Injury*, 13, 191-203.



## APPENDIX B – Discourse Error Analysis

<i>Error Type</i>	Codes	Level of Identification:	Operational Definition	Example
Superordinate substitution	Tier=%wor \$sup=(word)	Word	The substitution of a target word with its superordinate category, which conveys insufficient information for the content of the story.	e.g., “animal” for “dog,” “place” for “farm,” “thing” for “plant,” etc.
Word finding difficulties: Phonemic Paraphasia Formal Paraphasia Semantic Paraphasia Unrelated Paraphasia Unretrieved word	Tier=%wor \$par=phon=(word) \$par=sem=(word) \$par=form=(word) \$par=unr=(word) \$uw=(word)	Word	<u>Phonemic paraphasia</u> : The substitution of a target word with a non-word that is phonemically related to the target word (Davis, 2000).  <u>Formal paraphasia</u> : The substitution of a target word with a word that is phonemically related to the target word (Davis, 2000).  <u>Semantic paraphasia</u> : The substitution of a target word with a word that is related in meaning (Davis, 2000).  <u>Unrelated paraphasia</u> : The substitution of a target word with a word that has no apparent relation to the target (Davis, 2000).  <u>Unretrieved word</u> : The search for a specific vocabulary word that does not result in the retrieval of that word; an obvious and unsuccessful verbal search for a word.	Phonemic paraphasia: “fick” for “pick.”  Formal paraphasia: “kick” for “pick.”  Semantic paraphasia: “cat” for “dog.”  Unrelated paraphasia: “horse” for “spoon;” “log” for “cat.”  Unretrieved word: “I can’t remember, what is it, I can’t remember, etc.”
Perseveration	Tier=%phr \$per	Word	The repetition of a particular word that is NOT due to dysfluency. Perseveration occurs when the participant inappropriately continues to maintain one line of thinking without moving on (Frattali & Grafman, in press).	e.g., “Old MacDonald started to grow vegetables. And he grew vegetables. Then there were more vegetables.”
Violation of cohesive links: anaphora conjunction	Tier=%wor \$cos=anaph=(word) \$cos=conj=(word)	Word	<u>Violation of anaphora</u> : The use of nonreferential pronouns. If the transcriber cannot refer the pronoun back to its reference, apply this code (Ewing-Cobbs, Brookshire, Scott, & Fletcher, 1998).	Anaphora: e.g., “Keith drove to London yesterday. <u>It</u> kept breaking down.”

			<u>Violation of conjunction cohesion</u> : An error in the type of conjunction tie used to link utterances (Strong, 1998).	<u>Conjunction ties</u> : e.g., “The doctor told him to go home <i>so</i> he didn’t.” (The conjunction “so” was used erroneously for “but”).
Embellishment	Tier=%phr \$emb	Phrase/ Sentence	Exaggeration of story content (Frattali & Grafman, in press).	e.g., “The carrots were growing through the apartment every which way – up through the floor, down through the ceiling, through the side walls of the house.”
Confabulation	Tier=%phr \$con	Phrase/ Sentence	Fabrication of story content (Frattali & Grafman, in press).	e.g., “Mr. and Mrs. Old MacDonald got a divorce, and Old MacDonald moved back to the farmhouse.”
Perseveration	Tier=%phr \$per	Phrase/ Sentence	The repetition of a particular phrase, sentence, or thought that is NOT due to dysfluency.  The repetition does not have to be consecutive or verbatim.  Perseveration occurs when the participant inappropriately continues to maintain one line of thinking without moving on (Frattali & Grafman, in press).	e.g., “Mr. and Mrs. Old MacDonald started living in the city. They moved to the city. They came to live in the city. That’s where they were.”
Topic Stray	Tier=%phr \$str	Phrase/ Sentence	An utterance that is tangential to the topic of the story, including the addition of personal statements (Frattali & Grafman, in press).	e.g., The story is about Old MacDonald moving to an apartment in the city and starting a farm, and the participant begins to bring in personal information, i.e., “when I used to work on a farm...”
Faulty inference: predictive backward coherence	Tier=%phr \$inf=pre \$inf=bck \$inf=coh	Phrase/ Sentence	<u>Predictive</u> : An error in forward inferencing that reflects the readers’ inability to anticipate and/or predict a future event (Van Den Broek, 1994).  <u>Backward</u> : Faulty causal explanations connecting an event to its antecedents (Van Den Broek, 1994).	Predictive: e.g., “The apartment building owner would be pleased with the farm having taken over the apartment house.”  Backward: e.g., “So Old MacDonald decided to sell his produce in a stand because all of the other tenants were thrilled about what he

			<u>Coherence</u> : Faulty causal explanations for the events in the story, along with the inability to resolve ambiguities in reference to the event (Van Den Broek, 1994).	was doing.” <u>Coherence</u> : e.g., “They were all so unhappy that they decided to open a fruit and vegetable stand, to get rid of all of the fruits and vegetables in the apartment house.”
Faulty temporal ordering of events	Tier=%phr \$seq	Phrase/ Sentence	Misordering of temporal sequences of events that reflect an incomplete understanding or misunderstanding of the story (Frattali & Grafman, in press).	e.g., “First, they opened a fruit and vegetable stand. Then, the building owner got angry for having set up a farm in the apartment building.”
Difficulty interpreting gist	Tier=%glo \$gis	Thematic/Global	Partial or narrow ability or inability to capture theme or gist of the overall story (Frattali & Grafman, in press).	e.g., The inability to integrate the encounter with the building owner and subsequent creation of fruit and vegetable stand into story.

## Appendix References

Davis, A.G. (2000). *Aphasiology: Disorders and clinical practice*. Needham Heights, MA: Allyn & Bacon.

Ewing-Cobbs, L., Brookshire, B., Scott, M.A., & Fletcher, J. (1998). Children’s narrative following traumatic brain injury: Linguistic structure, cohesion, and thematic recall. *Brain and Language*, 61, 395-419.

Frattali, C. & Grafman, J. (2003; in press). Language and discourse deficits following prefrontal cortex damage. In L. L. LaPointe (Ed.), *Aphasia and related neurogenic language disorders*. New York: Thieme.

Strong, C.J. (1998). The Strong Narrative Assessment Procedure. Eau Claire, WI: Thinking Publications.

Van Den Broek, P. (1994). Comprehension and memory of narrative texts. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539-588). San Diego, CA: Academic Press.

## APPENDIX C – Pilot Questionnaire

DOB \_\_\_\_\_ Highest Level of Education obtained \_\_\_\_\_

Gender (circle one) M F Cultural Background (e.g., ESL) \_\_\_\_\_

1. (Circle one) Was the story:

*Easy* to understand *Moderately Easy* to understand *Difficult* to understand *Very Difficult* to understand?

2. Were there any pictures that were particularly difficult to process? YES NO

If yes, which ones and why?

---

---

---

3. Did you feel like you had enough time to process each frame of the story?

YES NO N/A

4. Were the comprehension questions clearly worded? YES NO

If not, which questions could be more clearly worded?

---

---

---

5. Did the comprehension questions fairly tap the information presented in the story?

YES NO

If not, which questions could be improved upon?

---

---

---

6. Did you feel like you had enough time to answer each comprehension question?

YES NO N/A

7. Do you have any other comments?

## **APPENDIX D – Compilation of Pilot Questionnaire Comments**

Q2: Were there any pictures that were particularly difficult to process? If yes, which ones and why?

- The picture involving the fountain, tenants, and several children (1 participant)
- Pictures towards the beginning of the story/overly detailed (7 participants)

Q5: Did the comprehension questions fairly tap the information presented in the story? If not, which questions could be improved upon?

- The story was so detailed that it was difficult to figure out what the most salient information was (1 participant)
- Questions too specific – don't catch them when reading for the "gist" (3 participants)

Q7: Do you have any other comments?

- Interesting story (2 participants)
- Absurdity of story was distracting (1 participant)
- It was hard to tell the people apart (1 participant)
- Was confused at beginning; got easier to understand as story progressed (2 participants)

# APPENDIX E – Instructions/Pictorial Stimuli/Comprehension Questions

## Instructions

This task helps me understand how people produce stories. You will see a series of pictures that tell a story. After viewing and understanding each picture, press the SPACE bar to continue.

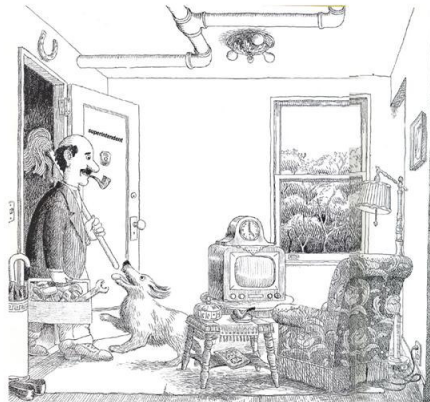
If you have not responded within 20 seconds, the computer will move to the next picture. Some of the frames will be detailed, so take your time and look carefully.

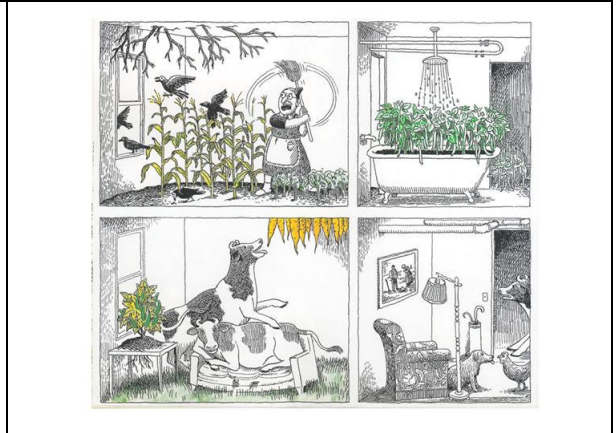
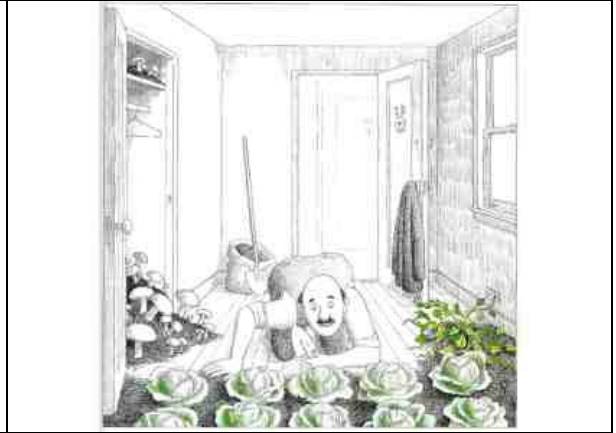
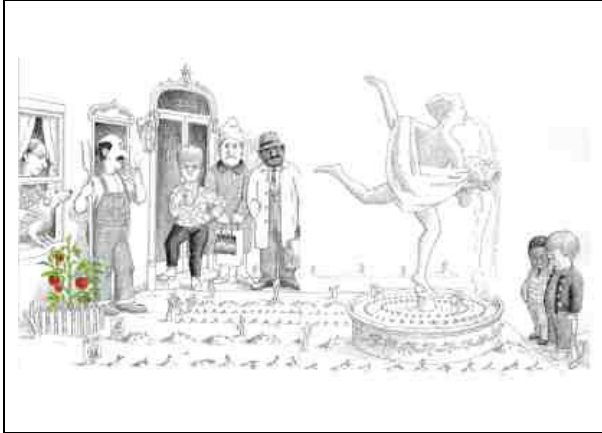
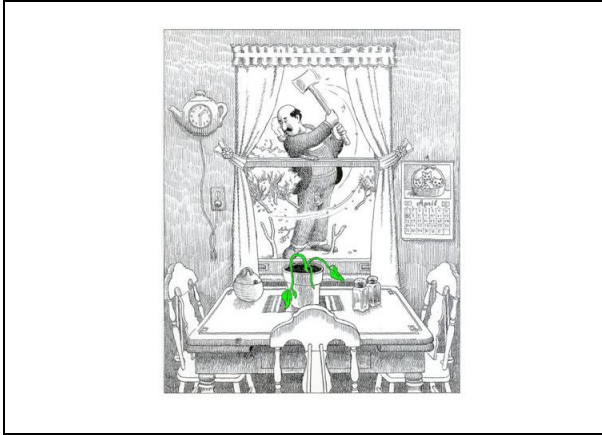
After you have finished viewing the pictures, I want you to, "Tell me the story, from beginning to end." Your telling of the story will be video-recorded.

Later on, I will ask you again to tell me about this story. At that time, you will be asked to respond to a set of questions about the story. So I want you to remember as much as you can about it.

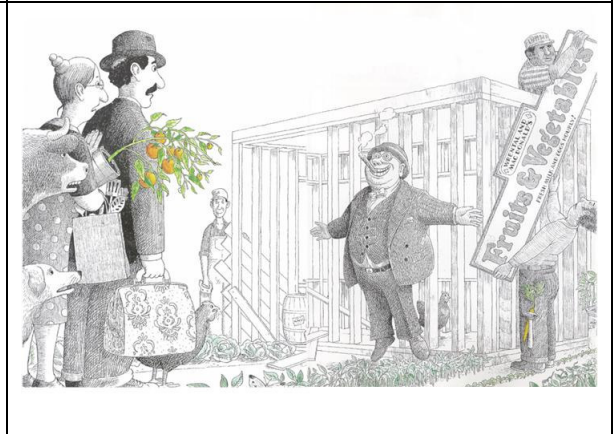
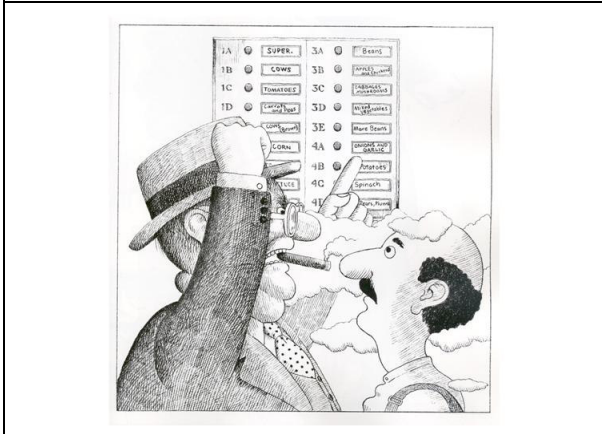
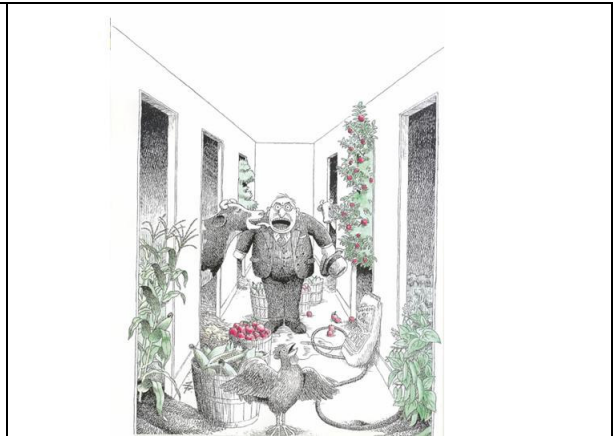
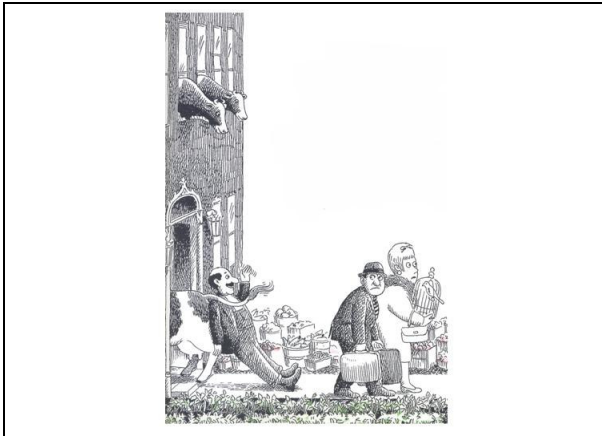
## Old MacDonald Had an Apartment House

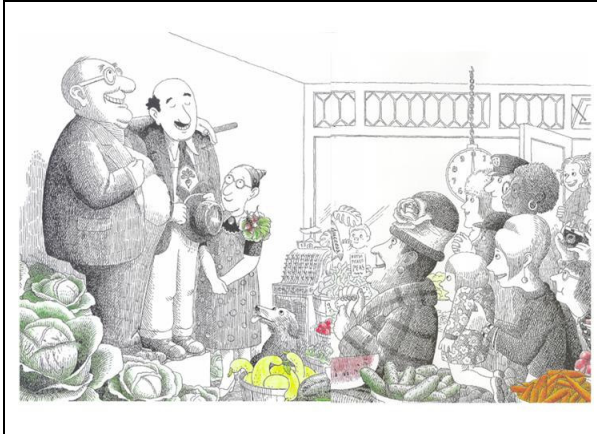
Modified from Barrett, J., & Barrett, R. (1969). Old MacDonald had an Apartment House. New York: Atheneum.











**Tell me the story from beginning to end.  
Be as specific as possible.**

<p>Do you remember the story you viewed earlier?</p> <p>I want you to tell me the story again from beginning to end. Be as specific as possible.</p>	<p>Now I want you to respond to a set of questions about the story.</p> <p>To get the questions, press the space bar. You will respond to a set of multiple-choice questions, and the responses will be A, B, C, or D.</p>
<p>The questions will appear on the computer screen, one at a time. To select your answer, use your index finger to press the labeled key corresponding to your answer.</p>	<p>Both accuracy and speed are important. I want you to respond quickly to the questions, but not so quickly that you make mistakes.</p> <p>If you have not responded within 30 seconds, the computer will go on to the next question. If you are not sure of an answer, go ahead and guess.</p>
<p style="text-align: center;"><b>Questions</b></p> <p>Press any key to answer the questions</p>	<p style="text-align: center;"><b>What was Old MacDonald's job?</b></p> <p>A. Lawyer B. Accountant C. Doorman D. Superintendent</p>

<p>Which two vegetables did Old MacDonald grow?</p> <ul style="list-style-type: none"><li>A. Apples and carrots</li><li>B. Carrots and lettuce</li><li>C. Corn and onions</li><li>D. Oranges and melons</li></ul>	<p>What person who smoked a cigar came to see Old MacDonald?</p> <ul style="list-style-type: none"><li>A. His wife</li><li>B. The landlord</li><li>C. A construction worker</li><li>D. A manager</li></ul>
<p>What first interested Old MacDonald in gardening?</p> <ul style="list-style-type: none"><li>A. A tree outside</li><li>B. His wife's tomato plant</li><li>C. His friend's garden</li><li>D. The bushes outside his window</li></ul>	<p>Why did Old MacDonald chop down the bushes at the beginning of the story?</p> <ul style="list-style-type: none"><li>A. To give the tomato plant more sun</li><li>B. Because the bushes were dying</li><li>C. To give him something to do</li><li>D. To plant them in his house</li></ul>
<p>How did Old MacDonald feel about changing the apartment building to a farm?</p> <ul style="list-style-type: none"><li>A. Sad</li><li>B. Angry</li><li>C. Embarrassed</li><li>D. Excited</li></ul>	<p>Besides cows, what other animals did the MacDonalds begin to raise?</p> <ul style="list-style-type: none"><li>A. Sheep</li><li>B. Horses</li><li>C. Chickens</li><li>D. Ducks</li></ul>

<p>What did Old MacDonald use to chop down the bushes?</p> <p>A. Shears B. An axe C. A saw D. A sledge hammer</p>	<p>Why did the landlord change his mind about making the MacDonalds move out?</p> <p>A. Because he could make a good profit selling produce B. Because he liked vegetables C. Because Old MacDonald had nowhere else to go D. Because he was given some of the good produce</p>
<p>What was the main point of this story?</p> <p>A. An apartment is a good place to grow vegetables B. Following your dreams pays off in the end C. It is easy to turn an apartment building into a farm D. It's okay to ask tenants to move out for a good cause</p>	<p>What was the last thing that happened in the story?</p> <p>A. Old MacDonald began to grow vegetables in the apartment building B. The tenants of the building got very angry C. The MacDonalds began selling their produce at a fruits and vegetables stand D. Old MacDonald chopped down the bushes in front of his apartment</p>
<p>What was the result of Old MacDonald planting carrots in the apartment?</p> <p>A. He was able to feed the poor B. The tenants below had to pull carrots from their ceiling C. He wanted to increase his vitamin intake D. He didn't have to buy food for his animals</p>	<p>Why was the landlord angry with Old MacDonald?</p> <p>A. Old MacDonald made too much noise B. Old MacDonald caused the tenants to leave C. A policeman complained to the landlord about Old MacDonald D. The landlord did not like pets</p>

## **APPENDIX F – Reliability Training Manual**

### **A PILOT STUDY TO DEVELOP A DISCOURSE ANALYSIS CODES SPECIFIC TO PREFRONTAL DYSFUNCTION**

Master's Thesis conducted under NIH protocol # 00-CC-0096: Investigations in  
Discourse Processes

Nan Ratner, Ph.D., Chair, Department of Hearing and Speech Sciences

#### **Reliability Training Procedures**

Inbal Eshel  
March 2004  
[ieshel@hesp.umd.edu](mailto:ieshel@hesp.umd.edu)  
(H) 301-986-1512  
(C) 301-502-1325

In participating in this project, you will be serving as a judge for the reliability of code assignments that I have created for my discourse analysis project, which will become part of my Master's Thesis. The codes are intended to capture the unique discourse deficits that present following prefrontal cortex damage.

Thank you so much for your participation. It is critical to this research project!

Thanks again,

Inbal

# A PILOT STUDY TO DEVELOP DISCOURSE ANALYSIS CODES SPECIFIC TO PREFRONTAL DYSFUNCTION

## Preliminary Project Summary

General descriptions of linguistic deficits yielded from conventional language tests cannot adequately characterize the discourse deficits of prefrontal cortex damaged (PFCD) patients. No system currently exists to analyze the unique discourse production deficits of the PFCD population. This pilot study sets forth the development of a discourse analysis system designed to capture the “nonaphasic” but characteristic discourse deficits that could present following PFCD. The study will utilize an analysis system based on narrative sample elicitation to investigate between-group differences in two study populations: patients with left, right, or bi-frontal PFCD and age and education-matched healthy comparison group participants. A primary focus will be the coding and comparison of narrative discourse features derived from transcripts on indices of content units, thematic units, and story grammar categories, in addition to a discourse error analysis. Our goal is to discover discourse profiles that will differentiate between the patient population and the healthy comparison group. Our findings will be of value in contributing to and advancing measurements specific to the behavioral manifestations of the PFCD population.

## PROCEDURE:

### Getting Started:

#### **BEFORE YOU BEGIN:**

Familiarize yourself with the story. Review it a minimum of 3 times to make sure that you have a good grasp of the content, the themes, the sequence of the story, and the characters.

#### General comments:

- Attached you'll find the transcripts of three subjects. Each subject viewed the story one time on the computer. Each subject told the story immediately after viewing it (story tell), and then 30 minutes later (story retell).
- You will be analyzing both the story tells and story retells of the three subjects. Therefore, there are two transcripts for each subject – the tell and retell.
- You will be entering some of the codes into an excel spreadsheet, and other codes into the transcript itself. More details to come....

Download a free program (CLAN) onto your computer in order to complete part of the coding.

#### HOW TO DOWNLOAD CLAN TO YOUR COMPUTER:

- 1) Go to <http://childes.psy.cmu.edu>
- 2) Go to Programs and Data
- 3) Download CLAN programs for appropriate operating system (e.g. Windows, Mac, etc.)
- 4) Download across the "OldCode" line – either CLANWIN or CLAN9 (depending on Mac or PC computer).
- 5) Install CLAN as directed.
- 6) It will install as CLAN. You can find it by going into your programs menu, and looking under "CLAN."
- 7) When you click on and open the program, it will open into the Command Window in front. You don't have to worry about this window – just close it to access the transcript window. If you have a macintosh, you may need to go to file and press "open" to get a transcript window.
- 8) Since you'll be using transcripts that have already been coded, just go to "open" and open the file that I will provide.



**Coding:**

There are four sets of codes that need to be entered for each participant:

- 1) Analysis by Content Unit (a frame-by-frame analysis to see which of the 16 frames the participant described in their story. Keep in mind that no participant will say the exact same utterance that is listed – make sure to be flexible.)
- 2) Analysis by Thematic Unit (a macro-level analysis to determine whether the participants captured the overall gist of the story, which has been broken down into five parts. Again, remember not to determine the presence or absence of a thematic unit verbatim – be flexible).
- 3) Analysis by Story Grammar features (a frame-by-frame analysis to see whether their descriptions match up with the story grammar codes associated with the frame).
- 4) Discourse error analysis (an utterance-level analysis in which various errors are coded at three levels: the word level, the phrase/sentence level, and at a global level).

**Order of Coding:**

Code in the above order – i.e., first content units, then thematic units, then story grammar, and finally discourse error analysis.

**If ever in doubt....**

Look back at the story to make sure.

### Ground rules for Content Unit Analysis:

- Make sure that you've read the story-tell/retell completely at least once.
- Look at the transcript utterance by utterance, and compare it to the frame by frame descriptions (see below).
- Determine whether or not each frame of the story was adequately covered in the story-tell and retell. Do not expect the stories of the participants to match the frame descriptions verbatim – just try to determine whether or not the information is adequately mentioned.
- Score each frame as either present (1) or absent (0) in the Excel file.

### Frame number and description

1. Old MacD and wife in front of apartment building, they move to the city.
2. Old MacD is superintendent of apartment building.
3. Wife is sad and holding dying plant that is not getting enough light.
4. Old MacD cuts down bushes in front of window to let the sunshine in.
5. Plant now has light and is thriving.
6. Old MacD begins planting seeds outside the building. Tenants are upset.
7. Old MacD begins growing vegetables inside the apartment building.
8. Carrots grow through the ceiling of one apartment; apartment dwellers are angry.
9. Fruits and vegetables are growing inside the apartments.
10. More tenants are forced to move out.
11. Building owner furious when sees that the building has been converted into a "farm."
12. Building owner gets angry at Old MacD for converting the apartment building into a farm.
13. Sad Old MacD and wife packing to leave.
14. Owner tries to decide what action to take because the plants are thriving.
15. Owner standing in front of construction of a fruit and vegetable stand. Old MacD and wife in shock.
16. Happy owner, Old MacD, and wife in front of fruit and vegetable stand; many customers.

### Ground rules for Thematic Unit Analysis:

- Make sure that you've read the story-tell/retell through completely at least once.
- Look at the transcript on a more holistic level, and compare it to the descriptions of the five thematic units (see below).
- Determine whether or not each thematic unit was adequately covered in the story-tell and retell. Do not expect the stories of the participants to match the thematic unit descriptions verbatim – just try to determine whether or not they are adequately covered.
- Score each frame as either present (1) or absent (0) in the Excel file.

### Thematic Unit and Description

1. Love of farming spurs Old MacDonald (the superintendent of an apartment building) to start growing produce in the building.
2. The apartment “farm” thrives.
3. Old MacDonald forces the apartment tenants to leave, to accommodate farm.
4. The building owner is angry and almost evicts Old MacDonald and his wife.
5. The building owner thinks of a “win-win” situation; let the Old MacDonald’s stay and open a fruit and vegetable stand in which to sell their produce.

Ground rules for Story Grammar Category Analysis: (See “Story Grammar – a quick lesson” for more information).

- I have gone through the story, identifying which story grammar (SG) category corresponds with each frame (see below).
- Look at the transcript utterance by utterance, and compare it to the frame by frame descriptions (see below).
- Determine whether or not each frame of the story was adequately covered in the story-tell and retell, and second whether or not they correspond with the set story grammar categories. Try to think about the intent of the story grammar category as described under “Story Grammar – a quick lesson”.
- Score each frame as either containing the specified story grammar category (1) or not (0) in the Excel file.

#### STORY GRAMMAR CATEGORIES:

SG:

#### Frame number and descriptor

1. Old MacD and wife in front of apt. bldg., they move to the city	S
2. Old MacD is sup't of bldg.	S
3. Wife is sad and holding dying plant that is not getting enough light.	IE
4. Old MacD cuts down bushes in front of window to let the sunshine in.	A
5. Plant now has light and is thriving.	O
6. Old MacD begins planting seeds outside the bldg. Tenants are upset.	G, O
7. Old MacD begins growing vegetables inside the apartment building.	A
8. Carrots grow through ceiling of one apartment; apartment dwellers are angry.	O
9. Fruits and vegetables are growing inside the apartments.	O
10. More tenants are forced to move out.	O
11. Building owner furious when sees that the building has been converted into a "farm."	O
12. Building owner gets angry at Old MacD for converting the apt. building into farm.	R
13. Sad Old MacD and wife packing to leave.	R
14. Owner tries to decide what action to take because the plants are thriving.	G
15. Owner standing in front of construction of a fruit and vegetable stand. Old MacD and wife in shock.	O
16. Happy owner, Old MacD, and wife in fruit and vegetable stand; many customers.	R

KEY -- story grammar categories:

S = Setting

IE = Initiating Event

IR = Internal Response

G = Goal/Plan

A = Attempt

O = Outcome

R = Reaction

Ground rules for Discourse Error Analysis: (For actual codes, see Discourse Error Analysis Table)

- Here you will be entering the data into CLAN. See Error Analysis Table for actual codes.
  - i. Enter both word and phrase/sentence-level codes underneath the utterance that they refer to. Do this by hitting “enter” after the utterance, and entering the applicable tier (e.g. “%wor or %phr) and code. Every utterance will begin with \*SU (for “subject) and then the number 1, 2, or 3.
  - ii. The one global-level code (“gist”) should be entered after the last code of the last utterance.
- Code errors in relation to how people typically relay stories. Be flexible in your approach to the transcript. Is the word/phrase/etc. acceptable, or is it really outside the norm of acceptability?
- Code at the three levels sequentially – e.g. code the entire story tell/retell at the word level first, then at the phrase/sentence level, then at the global level.
- The three levels should go one after another in the transcript.
  - i. E.g.,
    - \*SU1: Well Old MacDonald had an umbrella in his restaurant.
    - %wor: \$par=sem=umbrella
    - %phr: \$con
- Multiple coding is acceptable – any word, sentence, etc. can be coded more than once, if it fits into more than one category!
- If the participant self-corrects, don’t count it as an error (e.g., “The radish, I mean the carrot,” – let’s say that it would have been a semantic paraphasic error, but the patient self-corrects. It is not coded).
- If a word/phrase/etc. is repeated multiple times consecutively, think of it as a dysfluency, and only code the first instance in which the error occurs. (e.g. “has no pick um has no pick um” – if “pick” were a paraphasia, you would only code it one time).

**Story Grammar – a quick lesson:**

Story Grammar refers to the internal structure of stories.

It presumes that each story contains at least one of each of the following categories:

<i>Category</i>	<i>Definition</i>
Setting	<ul style="list-style-type: none"><li>• Backdrop for the story</li><li>• Describes characters, objects, geographical information, temporal information, etc.</li></ul>
Initiating event	<ul style="list-style-type: none"><li>• The event that sets the story in motion</li><li>• Obstacle, problem, or complication that causes a response from the protagonist(s)</li></ul>
Internal response	<ul style="list-style-type: none"><li>• The reaction of the protagonist/character to the initiating event</li></ul>
Goal/Plan	<ul style="list-style-type: none"><li>• The establishment of a goal, resulting from the internal response</li><li>• Statements about how the character(s) might overcome the obstacle, solve the problem, or deal with the complication</li></ul>
Attempt	<ul style="list-style-type: none"><li>• The various ways in which the protagonist tries to reach the goal</li><li>• Actions taken by the protagonist(s)</li></ul>
Outcome	<ul style="list-style-type: none"><li>• The result of the various attempts to reach the goal</li><li>• The attempt's aftermath, or repercussions of the attempts to overcome the obstacle, solve the problem, or deal with the complication</li></ul>
Reaction	<ul style="list-style-type: none"><li>• The reaction of the protagonist to the success or failure of the outcome</li><li>• Thoughts, emotions, and actions of the character(s) at the end of the story</li></ul>

Subject 1

Transcription and Coding Checklist:

- Coding: **STORY TELL**
  - i. Content Units (entered into Excel file) \_\_\_\_\_
  - ii. Thematic Units (entered into Excel file) \_\_\_\_\_
  - iii. Story Grammar Categories (entered into Excel file) \_\_\_\_\_
  - iv. Discourse Error Analysis (entered into CLAN) \_\_\_\_\_
- Coding: **STORY RETELL**
  - i. Content Units (entered into Excel file) \_\_\_\_\_
  - ii. Thematic Units (entered into Excel file) \_\_\_\_\_
  - iii. Story Grammar Categories (entered into Excel file) \_\_\_\_\_
  - iv. Discourse Error Analysis (entered into CLAN) \_\_\_\_\_
- PRINT a copy

Notes:

Subject 2

Transcription and Coding Checklist:

- Coding: **STORY TELL**
  - i. Content Units (entered into Excel file) \_\_\_\_\_
  - ii. Thematic Units (entered into Excel file) \_\_\_\_\_
  - iii. Story Grammar Categories (entered into Excel file) \_\_\_\_\_
  - iv. Discourse Error Analysis (entered into CLAN) \_\_\_\_\_
- Coding: **STORY RETELL**
  - i. Content Units (entered into Excel file) \_\_\_\_\_
  - ii. Thematic Units (entered into Excel file) \_\_\_\_\_
  - iii. Story Grammar Categories (entered into Excel file) \_\_\_\_\_
  - iv. Discourse Error Analysis (entered into CLAN) \_\_\_\_\_
- PRINT a copy

Notes:

Subject 3

- Coding: **STORY TELL**
  - i. Content Units (entered into Excel file) \_\_\_\_\_
  - ii. Thematic Units (entered into Excel file) \_\_\_\_\_
  - iii. Story Grammar Categories (entered into Excel file) \_\_\_\_\_
  - iv. Discourse Error Analysis (entered into CLAN) \_\_\_\_\_
- Coding: **STORY RETELL**
  - i. Content Units (entered into Excel file) \_\_\_\_\_
  - ii. Thematic Units (entered into Excel file) \_\_\_\_\_
  - iii. Story Grammar Categories (entered into Excel file) \_\_\_\_\_
  - iv. Discourse Error Analysis (entered into CLAN) \_\_\_\_\_
- PRINT a copy

Notes:

## WORKS CITED

- Adams, C. (2002). Practitioner Review: The assessment of language pragmatics. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 43, 973-88.
- Alexander, M.P. (2002). Disorders of language after frontal lobe injury: Evidence for the neural mechanisms of assembling language. In D.T. Stuss, & R.T. Knight, (Eds.) *Principles of frontal lobe function*. (pp. 159-167). New York: Oxford University Press.
- Barrett, J., & Barrett, R. (1969). *Old MacDonald had an Apartment House*. New York: Atheneum.
- Beeman, M. (1993). Semantic processing in the right hemisphere may contribute to drawing inferences from discourse. *Brain and Language*, 44, 80-120.
- Binder, J.R., Frost, J.A., Hammeke, T.A, Cox, R.W., Rao, S.M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *Journal of Neuroscience*, 17, 353-62.
- Bliss, L.S., McCabe, A., & Miranda, A.E. (1998). Narrative Assessment Profile: Discourse analysis for school-age children. *Journal of Communication Disorders*, 31, 347-363.
- Brinton, B., & Fujiki, M. (1989). *Conversational management with language-impaired children: Pragmatic assessment and intervention*. Gaithersburg, MD: Aspen.



- Brookshire, R.H., & Nicholas, L.E. (1993). *Discourse Comprehension Test*.  
Arizona: Communication Skill Builders.
- Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge, MA: Cambridge  
University Press.
- Cannizzaro, M.S., & Coelho, C.A. (2002). Treatment of story grammar following  
traumatic brain injury. *Brain Injury, 16*, 1065-1073.
- Caplan, D. (1999). *Language: Structure, processing, and disorders*. Cambridge,  
MA: MIT Press.
- Clark, H.H. (1994). Discourse in production. In M.A. Gernsbacher (Ed.), *Handbook  
of psycholinguistics* (pp. 985-1021). San Diego, CA: Academic Press.
- Coelho, C.A. (2002). Story narratives of adults with closed head injury and non-  
brain-injured adults: Influence of socioeconomic status, elicitation task, and  
executive functioning. *Journal of Speech, Language, and Hearing Research,*  
*45*, 1232-1248.
- Craig, G., & Frattali, C.M. (2000). "Getting the gist" of stories: Effects of prefrontal  
cortex damage. *ASHA Leader, 5*, 219.
- Damico, J.S. (1985). Clinical discourse analysis: A functional approach to language  
assessment. In C. S. Simon (Ed.), *Communication Skills and Classroom  
Success: Assessment of language-learning disabled students* (pp. 165-204).  
San Diego, CA: College-Hill Press.
- Davis, A.G. (2000). *Aphasiology: Disorders and Clinical Practice*. Needham  
Heights, MA: Allyn & Bacon.

- Davis, A.G., O'Neil-Pirozzi, T.M., & Coon, M. (1997). Referential cohesion and logical coherence of narration after right hemisphere stroke. *Brain and Language, 56*, 183-210.
- Demb, J.B., Desmond, J.E., Wagner, A.D., Vaidya, C.J., Glover, G.H., & Gabrieli, J.D. (1995). Semantic encoding and retrieval in the left inferior prefrontal cortex: A functional MRI study of task difficulty and process specificity. *Journal of Neuroscience, 15*, 5870-8.
- Ewing-Cobbs, L., Brookshire, B., Scott, M.A., & Fletcher, J. (1998). Children's narrative following traumatic brain injury: linguistic structure, cohesion, and thematic recall. *Brain and Language, 61*, 395-419.
- Ferstl, E.C., Guthke, T., & Cramon, D.Y. (1999). Change of perspective in discourse comprehension: Encoding and retrieval processes after brain injury. *Brain and Language, 70*, 385-420.
- Fleiss, J.L. (1981) *Statistical methods for rates and proportions, 2<sup>nd</sup> edition.* (p. 218) New York: John Wiley & Sons.
- Frattali, C., & Grafman, J. (in press). Language and discourse deficits following prefrontal cortex damage. In L. L. LaPointe (Ed.), *Aphasia and related neurogenic language disorders.* New York: Thieme.
- Fuster, J.M. (1997). *The prefrontal cortex: Anatomy, physiology, and neuropsychology of the frontal lobe, 3<sup>rd</sup> edition.* Philadelphia, PA: Lippincott-Raven.
- Garrett, K. (1986). Telling tales: Narratives and learning disabled children. *Topics in Language Disorders, 6*, 44-56.

- Gernsbacher, M.A., & Varner, K.R. (1988). The multi-media comprehension battery. Eugene, University of Oregon.
- Glosser, G., & Deser, T. (1990). Patterns of discourse production among neurological patients with fluent language disorders. *Brain and Language*, 40, 67-88.
- Grafman, J. (2002). The structured event complex and the human prefrontal cortex. In D.T. Stuss & R.T. Knight (Eds.), *The frontal lobes* (pp. 292-310). New York: Oxford University Press.
- Gregg, N., & Hoy, C. (1989). Coherence: The comprehension and production abilities of college writers who are normally achieving, learning disabled, and underprepared. *Journal of Learning Disabilities*, 22, 370-2, 390.
- Halliday, M.A.K. (1994). *An introduction to functional grammar* (2<sup>nd</sup> ed.). London: Edward Arnold.
- Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- 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.
- Jurica, P.J., Leitten, C.L., & Mattis, S. (2001). *Dementia Rating Scale-2*. Florida: Psychological Assessment Resources, Inc.
- Kapur, S., Rose, R., Liddle, P.F., Zipursky, R.B., Brown, G.M., Stuss, D., Houle, S., & Tulving, E. (1994). The role of the left prefrontal cortex in verbal processing: Semantic processing or willed action? *Neuroreport*, 5, 2193-6.

- Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., & Fox, P.T. (2000). Automated talairach atlas labels for functional brain mapping. *Human Brain Mapping, 10*, 120-131.
- Liles, B.Z. (1985). Cohesion in the narratives of normal and language disordered children. *Journal of Speech and Hearing Research, 28*, 123-133.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Makale, M., Solomon, J., Patronas, N.J., Danek, A., Butman, J.A., Grafman, J. (2002). Quantification of brain lesions using interactive automated software. *Behavior Research Methods, Instruments, & Computers, 34*, 6-18.
- Mayer, M. (1969). *Frog, where are you?* New York: Dial Books for Young Readers.
- McCord, J.S., & Haynes, W.O. (1988). Discourse errors in students with learning disabilities and their normally achieving peers: molar versus molecular views. *Journal of Learning Disabilities, 21*, 237-243.
- McDonald, S. (1993). Viewing the brain sideways? Frontal versus right hemisphere explanations of non-aphasic language disorders. *Aphasiology, 7*, 535-549.
- McDonald, S. (2000). Exploring the cognitive basis of right-hemisphere pragmatic language disorders. *Brain and Language, 75*, 82-107.
- Menn, L., Ramsberger, G., & Helm-Estabrooks, N. (1994). A linguistic communication measure for aphasic narratives. *Aphasiology, 8*, 343-359.

- Mentis, M., & Prutting, C.A. (1987). Cohesion in the discourse of normal and head-injured adults. *Journal of Speech and Hearing Research, 30*, 88-98.
- Mentis, M., & Prutting, C.A. (1991). Analysis of topic as illustrated in a head-injured and a normal adult. *Journal of Speech and Hearing Research, 34*, 583-595.
- Miller, J., & Chapman, R. (1985). *Systematic analysis of language transcripts*.  
Madison, WI: University of Wisconsin.
- Montague, M., Maddux, C.D., & Dereshiwsky, M.I. (1990). Story grammar and comprehension and production of narrative prose by students with learning disabilities. *Journal of Learning Disabilities, 23*, 190-197.
- Morris-Friehe, M.J., & Sanger, D.D. (1992). Language samples using three story elicitation tasks and maturation effects. *Journal of Communication Disorders, 25*, 107-24.
- Moya, K.L., Benowitz, L.I., Levine, D.N., & Finklestein, S.P. (1986). Covariant defects in visuospatial abilities and recall of verbal narrative after right hemisphere stroke. *Cortex, 22*, 381-397.
- Poldrack, R.A., Wagner, A.D., Prull, M.W., Desmond, J.E., Glover, G.H., & Gabrieli, J.D. (1999). Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. *Neuroimage, 10*, 15-35.
- Portney, L.G., & Watkins, M.P. (2000). *Foundations of clinical research: Applications to practice, 2<sup>nd</sup> edition*. New Jersey: Prentice-Hall, Inc.

- Rochon, E., Saffran, E.M., Berndt, R.S., & Schwartz, M.F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language, 72*, 193-218.
- Roth, F.P., & Spekman, N.J. (1986). Narrative discourse: Spontaneously generated stories of learning-disabled and normally achieving students. *Journal of Speech and Hearing Disorders, 51*, 008-023.
- Roth, F.P., & Spekman, N.J. (1994). Oral story production in adults with learning disabilities. In R.L. Bloom, L.K. Obler, S. De Santi, & J.S. Erlich. (Eds.) *Discourse Analysis and Applications: Studies in Adult Clinical Populations* (pp. 131-147). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Saffran, E.M., Berndt, R.S., & Schwartz, M.F. (1980). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language, 37*, 440-479.
- Shewan, C. (1979). Auditory Comprehension Test for Sentences. Chicago, IL: Biolinguistics Clinical Institutes.
- Sirigu, A., Zalla, T., Pilon, B., Grafman, J., Dubois, B., & Agid, Y. (1995). Planning and script analysis following prefrontal lobe lesions. *Annals of the New York Academy of Sciences, 769*, 277-288.
- Snow, P.C., Douglas, J.M., & Ponsford, J.L. (1999). Narrative discourse following severe traumatic brain injury: A longitudinal follow-up. *Aphasiology, 13*, 529-551.
- Strong, C.J. (1998). The Strong Narrative Assessment Procedure. Eau Claire, WI: Thinking Publications.

- Talairach, J., & Tournoux, P. (1988). *Co-Planar stereotaxic atlas of the human brain*. Thieme: New York.
- Togher, L. (2001). Discourse sampling in the 21<sup>st</sup> century. *Journal of Communication Disorders, 34*, 131-150.
- Togher, L., & Hand, L. (1999). The macrostructure of the interview: Are traumatic brain injury interactions structured differently to control interactions? *Aphasiology, 13*, 709-723.
- Tompkins, C.A. (1995). *Right hemisphere communication disorders: Theory and management*. San Diego, CA: Singular Publishing Group, Inc.
- Tompkins, C.A., Baumgaertner, A., Lehman, M.T., & Fossett, T.R.D. (1997). Suppression and discourse comprehension in right brain-damaged adults. A preliminary report. *Aphasiology, 11*, 505-519.
- Tucker, F.M., & Hanlon, R.E. (1998). Effects of mild traumatic brain injury on narrative discourse production. *Brain Injury, 12*, 783-792.
- Ulatowska, H.K., Freedman-Stern, R., Doyle, A.W., & Macaluso-Haynes, S. (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 Den Broek, P. (1994). Comprehension and memory of narrative texts. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539-588). San Diego, CA: Academic Press.

- Wapner, W., Hamby, S., & Gardner, H. (1981). The role of the right hemisphere in the apprehension of complex linguistic materials. *Brain and Language, 14*, 15-32.
- Wechsler, D. (1997). Wechsler Memory Scale – Third Edition. The Psychological Corporation.
- Wood, J.N., & Grafman, J. (2003). Human prefrontal cortex: processing and representational perspectives. *Nature Reviews Neuroscience, 4*, 139-147.
- Wozniak, R.J., Coelho, C.A., Duffy, R.J., & Liles, B.Z. (1999). Intonation unit analysis of conversational discourse in closed head injury. *Brain Injury, 13*, 191-203.
- Zalla, T., Phipps, M., & Grafman, J. (2002). Story processing in patients with damage to the prefrontal cortex. *Cortex, 38*, 215-31.



