View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by The Aphasiology Archive

Introduction

Discourse production is a metalinguistic skill that is crucial for daily communicative interactions. Despite largely intact microlinguistic abilities, patients with prefrontal cortex damage (PFCD) often present with impaired discourse-level processes (Ferstl, Guthke, & Cramon, 1999) that can profoundly impact life routines and community re-integration. Narrative discourse such as storytelling is especially vulnerable in this population (Tucker & Hanlon, 1998). It is well known that conventional language tests do not adequately capture these discourse impairments. There is currently a dearth of discourse analyses procedures and empirical data on patients with PFCD.

Descriptive observations of narrative discourse production such as storytelling reveal that patients with PFCD may have difficulty recalling narrative components of a story, processing and drawing inferences, and appreciating a story's thematic aspects or gist. Other discourse deficits may include confabulation, embellishment, topic stray, ambiguous statements, faulty anaphoric reference and links, loss of moralistic meaning, and faulty temporal sequencing of events (Ferstl et al., 1999; Frattali & Grafman, 2005; Zalla, Phipps, & Grafman, 2002). Discourse procedures developed for other populations may not tap into the unique aspects of PFCD discourse.

This research outlines the development and initial validation of a computer-assisted discourse analysis system designed to identify the characteristic narrative discourse production deficits that may present following PFCD. We hypothesized that the PFCD group and matched healthy controls would differ on measures of lexical efficiency, content units, thematic units, story grammar features, and qualitative discourse errors in both immediate and delayed narrations, with poorer overall performance on the delayed condition, an increasingly difficult task. Inter-coder reliability was assessed to validate the clinical usefulness of this procedure.

Method

Participants

Ten participants were included: five with left, right, or bilateral lesions confined to the PFC (3 males; Mean age=54 y [SD=12.64]; Mean education=18y [SD=2.68]), and five age-, gender- and education- pairwise matched non-brain damaged healthy controls (3 males; Mean age=52 [SD=13.77]; Mean education=17 [SD=2.55]). All participants were right handed, native English speakers recruited from the participant database at the National Institutes of Health.

Procedure

Narratives were elicited following an adapted procedure from Gernsbacher and Varner (1988) using a pictorial story (modified from *Old MacDonald Had an Apartment House*, Barrett & Barrett, 1969) that was presented frame-by-frame on a PC running the SuperlabPro™ 1.74 software program. Data from fourteen healthy young volunteers were used to refine the narrative elicitation stimuli and procedure prior to use for this study.

After participants viewed the story frames in a self-paced manner, three responses were elicited: 1) immediate narration of the story with instructions to include as much detail as possible; 2) delayed narration after thirty minutes; 3) answers to thirteen multiple-choice comprehension questions on SuperlabPro™ 1.74.

Data Analysis

The first author, blinded to group assignment, transcribed each audiorecording orthographically into the Computerized Language Analysis system (CLAN; MacWhinney, 2003), an automated program for quantitatively analyzing language transcripts. Lexical efficiency was calculated using CLAN with the following formula (adapted from Menn, Ramsberger, & Helm-Estabrooks, 1994): total words [excluding fillers]/total number of story content units. Each transcript was then analyzed on four levels, the first three of which were based on a template of the story created by the first author:

I. Content units: frame-by-frame analysis of story content (16 total frames)

II. Thematic units: macro-level analysis of story gist (5 total themes)

III. Story grammar: setting, initiating event, internal response, goal, attempt, outcome, and reactions (Van Den Broek, 1994)

IV. Discourse errors that are typically associated with PFCD (see Table 1 for list). Word-finding difficulties (unretrieved words, phonemic/formal/semantic/unrelated paraphasias, and superordinate substitutions) were also coded.

Coding reliability

This was obtained by two trained judges, each of whom rated a random sample of 30% of total transcripts (blinded to participant and group assignment). Discrepancies were resolved by majority consensus. Overall, inter-rater reliability comparing the codes assigned by each judge to those assigned by the other two was fair to good (Kappa statistic ranged from 0.32 to 0.737) for levels I through III. An Intraclass Correlation Coefficient was utilized to assess inter-rater reliability for level IV, and ranged from good to excellent (0.604 to 0.857). Measures that had poor inter-rater reliability (below an ICC or Kappa statistic of 0.3) were not analyzed further.

The 13 multiple-choice questions (each with responses A through D) were answered in a self-paced manner and scored directly on SuperlabPro^{TM} 1.74.

Results

The values obtained for each discourse measure in each group are provided in Table 1. Lexical efficiency analysis was significantly poorer for the PFCD group (Wilcoxon signed-rank test, p = .04) across both conditions. These results indicate that the PFCD group was significantly less efficient in their storytelling (higher numbers correspond with lower efficiency).

Although the stories of the PFCD group contained fewer content units, thematic units, and story grammar features than the healthy comparison group during immediate and delayed recall conditions, no statistically significant differences emerged based on Wilcoxon signed-rank tests. The stories of the PFCD group contained more discourse errors than the stories of the healthy comparison group in every category, with several statistically significant between-group differences at $p<.05$ (see Table 1). For example, on the immediate recall phase, there were ten uncoded discourse errors in the healthy comparison groups narratives, and only three uncoded discourse errors in the PFCD group.

There were no significant group differences in accuracy of comprehension questions (77% and 89% mean for PFCD and control group respectively; Mann-Whitney Test p = 0.280).

Discussion

This is the first study to pilot a set of quantitative and qualitative measures designed specifically to capture the discourse impairments of PFCD patients as described in the literature. Our results support the original predictions. First, consistent with findings from other populations (e.g. healthy older vs. young adults, Blake, 2006; Right Hemisphere Brain Damage, Mackenzie, Begg, Lees, & Brady, 1999), lexical efficiency clearly and consistently differentiated between the PFCD population and the comparison group across all levels of analysis. Although the analyses by content, thematic, and story grammar units did not statistically differentiate between the groups, the consistent trends, in which the PFCD group narratives contained less content, would likely reach statistical significance given a larger sample size. The discourse errors found in the narrative samples, some statistically significant, also clearly differentiated between the two groups. The occurrence of discourse errors in the comparison group is not surprising and has been reported in the literature (Armstrong, 2002). The delayed recall component appears to add value and sensitivity to the overall task composition because this task, when combined with immediate recall, increased the number of measures that were statistically significantly different. Our finding of variability in inter-rater reliability is consistent with the discourse analysis literature in healthy and disordered populations (Armstrong, 2002; Body & Perkins, 1998, respectively), and may limit clinical utility.

Future research should enhance the power of the analyses and the specificity and sensitivity of the tools. For example, as has been found in other populations (e.g. RHBD, Blake, 2006), based on its sensitivity, it is recommended that lexical efficiency and other consistently coded elements (i.e. anaphoric cohesion, embellishment) be included in future discourse analyses. Conversely, elements that did not appear in any transcripts may not be sensitive to the discourse deficits of PFCD patients (i.e. phonemic paraphasias, conjunction cohesion, faulty predictive inferencing), and may be eliminated from future studies.

The study of discourse production in the PFCD population is critical, as deficits in these areas can significantly impact successful social and community reintegration (e.g. Galski, Tompkins, & Johnston, 1998). Subsequent to refinement, replication, and normative data collection, the tools created here are intended to assist in the diagnosis of PFCD patients, provide a basis for therapeutic intervention, and track changes in discourse over time. Future direction includes hopes of advancing clinical measurement in other patient populations with discourse-level deficits, including but not limited to schizophrenia, epilepsy, learning disabilities, TBI, and RHBD.

Table 1. Narrative Discourse Analysis. Wilcoxon signed-rank test scores.

Analysis Level	Measure	Immediate Recall		Delayed Recall		Immediate & Delayed	
		Control	PFCD	Control	PFCD	Control	PFCD
Lexical Efficiency	Total words/ total # story content units	34.08 (13.4) *	70.21 (38.39) *	32.47 (11.93) *	60.17 (25.14) *	33.27 (11.99) *	65.19 (31.05) *
Content Units	Total story content units	10.40 (2.608)	7.0 (2.55)	9.60 (2.074)	7.2 (3.194)	20.0 (4.637)	14.2 (5.675)
Thematic Units	Total thematic units	3.40 (1.517)	2.4 (.894)	3.20 (1.304)	2.2 (1.095)	6.6 (2.793)	4.6 (1.949)
Story Grammar Categories	Total story grammar categories	10.40 (4.506)	7.20 (1.789)	9.20 (3.899)	8.0 (2.345)	19.6 (8.385)	15.2 (3.894)
Discourse Error Analysis	Total Word	2.0 (.707) *	12.20 (7.12) *	1.40 (1.14)	10.0 (6.44)	1.7 (.949) *	11.1 (6.5) *
	Total Phrase	1.40 (1.673)	8.0 (6.12)	1.0 (1.23)	6.8 (3.56)	1.20 (1.40) *	7.4 (4.77) *
	Semantic Paraphasias - W	1.20 (.447)	5.80 (5.93)	.6 (.89)	3.60 (4.28)	.9 (.738) *	4.7(5.01) *
	Phonemic Paraphasias - W	0	.20 (.45)	0	0	0	.1 (.32)
	Formal Paraphasias - W	0	.20 (.45)	0	.60 (.89)	0	.40 (.70)
	Superordinate Substitution - W	.20 (.45)	2.20 (1.79)	.40 (.55)	1.4 (1.95)	.3 (.48)	1.8 (1.8)
	Unrelated Word - W	.20 (.45)	.40 (.54)	0	.20 (.45)	.1 (.32)	.3 (.48)
	Unretrieved Word - W	0	.40 (.54)	0	0	0	.2 (.42)
	Perseveration - W	0	.20 (.45)	0	.80 (1.30)	0	.5 (.97)
	Anaphoric Cohesion - W	0.4 (.89)	2.8 (1.48)	.40 (.55) *	3.40 (2.19) *	.40(.70)*	3.1(1.80)*
	Conjunction Cohesion - W	0	0	0	0	0	0
	Embellishment - P	0	2.80 (2.28)	0 *	3.4 (2.70) *	0*	3.1(2.38)*
	Confabulation - P	.20 (.45)	1.0 (1.22)	0	1.6 (1.14)	0.1 (.32)*	1.3(1.16)*
	Perseveration - P	0	0	0	0	0	0
	Topic Stray - P	0	1.0 (2.24)	0	.6 (0.89)	0	.8(1.62)
	Faulty Predictive Inferencing - P	0	0	0	0	0	0
	Faulty Backward Inferencing - P	0	0.2 (.45)	0	0	0	.1(.32)
	Faulty Coherence Inferencing - P	1 (1.23)	2 (2.35)	1 (1.22)	1 (1.41)	1(1.16)	1.50(1.90)
	Faulty Temporal Sequencing - P	.20 (.45)	1.0 (1.22)	0	.2 (.45)	.1(.32)	.6(.97)

* Significant at p<.05

Mean (SD); W = word level; P = phrase level

Note: Similar results found with between-group paired t-tests (the square root of the number of errors in each category was taken prior to analysis to normalize the distribution).

References

Armstrong, E. (2002). Variation in the discourse of non-brain-damaged speakers on a clinical task. *Aphasiology, 16*(4/5/6), 647-658.

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

Blake, M.L. (2006). Clinical relevance of discourse characteristics after right hemisphere brain damage. *American Journal of Speech-Language Pathology, 15*, 255-267.

Body, R., & Perkins, M.R. (1998). Ecological validity in assessment of discourse in traumatic brain injury: ratings by clinicians and non-clinicians. *Brain Injury, 12*(11), 963-976.

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.

Frattali, C., & Grafman, J. (2005). Language and discourse deficits following prefrontal cortex damage. In L. L. LaPointe (Ed.), *Aphasia and Neurogenic Language Disorders* (Third edition, pp. 51-67). New York, NY: Thieme.

Galski, T., Tompkins, C., & Johnston, M.V. (1998). Competence in discourse as a measure of social integration and quality of life in persons with traumatic brain injury. *Brain Injury, 12*(9), 769-782.

Gernsbacher, M.A., & Varner, K.R. (1988). The multi-media comprehension battery. Eugene, University of Oregon.

Mackenzie, C., Begg, T., Lees, K.R., & Brady, M. (1999). The communication effects of right brain damage on the very old and the not so old. *Journal of Neurolinguistics, 12,* 79-93.

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.

Menn, L., Ramsberger, G., & Helm-Estabrooks, N. (1994). A linguistic communication measure for aphasic narratives. *Aphasiology, 8,* 343-359.

Tucker, F.M., & Hanlon, R.E. (1998). Effects of mild traumatic brain injury on narrative discourse production. *Brain Injury, 12,* 783-792.

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.

Zalla, T., Phipps, M., & Grafman, J. (2002). Story processing in patients with damage to the prefrontal cortex. *Cortex, 38,* 215-31.