Production at the discourse level entails the activation and interaction of multiple interconnected cognitive and linguistic subsystems. However, people with aphasia (PWA) typically present with language processing deficits that can significantly impact how successfully they communicate. Language deficits can impact the flow of ideas and make verbal output appear disjointed and tangential, with loose associations between discourse units. An important aspect, then, of aphasia assessment is the analysis of discourse production. Discourse production has the advantages that it occurs naturally and is a common form of communication that offers an opportunity to observe complex cognitive/linguistic behaviors, and allows clinicians and researchers to conduct a wide variety of analyses to understand the nature of cognitive-communicative deficits.

Discourse coherence is one of the macro-linguistic properties that have attracted considerable interest in recent years. Coherence refers to the conceptual organization of discourse and it can be subdivided into two types: global and local. This paper focuses on global coherence (GC), which reflects how discourse relates to a global theme (Glosser & Deser, 1990; Kintsch & van Dijk, 1978).

Relatively little is known about coherence in aphasia; review of the studies investigating coherence indicates a range of performance among PWA (e.g., Coelho & Flewellyn, 2003; Glosser & Deser, 1990). Findings may reflect differences in how coherence is conceptualized (Armstrong, 2000), or low measurement accuracy due to the complex and abstract nature of the construct and lack of measurement instruments that facilitate valid score interpretations. Even though GC has been measured in PWA using several measurement methods (i.e., rating scales, coherence violations, total counts; Christiansen, 1995; Coelho & Flewellyn, 2003; Glosser & Deser, 1990; Ulatowska, Olness, & Williams, 2004), the psychometric properties of such methods have not been systematically examined.

The first goal of this study is to explore the psychometric properties of three measures of GC in terms of external aspects of validity (convergent and divergent), as well as their reliability over repeated samplings of story-telling. The first measure is a 5-point rating scale that has been developed by Glosser and Deser (1990) and has been traditionally used to measure GC of aphasic discourse. The second measure is a recently developed 4-point rating scale that conceptualizes discourse coherence similarly to Glosser and Deser (1990) but has demonstrated stability across age groups and discourse types, and good inter- and intra-rater reliability (Koutsoftas, Wright, & Capilouto, 2009).

The third measure is based on latent semantic analysis (LSA) which is a cognitive computational model of human knowledge acquisition (for details see Landauer & Dumais, 1997; http://lsa.colorado.edu/). The performance of the LSA algorithm has been shown to approximate human cognitive-semantic relations with respect to areas such as language acquisition, episodic memory, semantic priming, semantic categorization, and the effects of text coherence on comprehension (Foltz, Kintsch, & Landauer, 1998; Howard, Jing, Addis, & Kahana, 2007; Landauer & Dumais, 1997). Recently, Elvevåg, Foltz, Weinberger, and Goldberg (2007) used LSA to quantify discourse coherence in a group of individuals with schizophrenia. Using LSA, Elvevåg et al. were able to localize where in sentence production incoherence occurs, predict levels of incoherence, and identify whether discourse "belonged" to a patient or participant with typical language skills.

Two additional goals of this study are (i) to investigate whether there are differences between PWA and neurologically intact individuals in terms of GC and, (ii) assess the magnitude of the relationship between GC and severity of aphasia.

METHOD

To date, fifteen adults with aphasia have completed the study protocol. Inclusion criteria were as follows: (1) monolingual, English speakers; (2) at least six months post onset of stroke; (3) single, left-hemisphere CVA, and (4) sufficient hearing and visual acuity as indicated by passing hearing and vision screenings. Aphasia presentation was confirmed through performance on the Western Aphasia Battery-Revised (WAB-R; Kertesz, 2007) and clinical judgment. Fifteen adults without neurological impairment (NI) served as the control group. Inclusion criteria were as follows: (1) sufficient hearing and visual acuity as indicated by passing hearing and vision screenings (2) reported negative history for cognitively deteriorating conditions such as Alzheimer's or Parkinson's, and (3) normal cognitive functioning as measured by the *Mini-Mental Status Examination* (MMSE; Folstein, Folstein, & McHugh, 1975). All study participants completed four different discourse tasks including the story telling task used for the present study. Participants told the stories depicted in two wordless pictures books: *Good Dog Carl* (1985) and *Picnic* (McCully, 1984).

The language samples have been orthographically transcribed, segmented into c-units, and subjected to coherence analyses using each of the three measures. First, coherence was calculated using a 4-point rating scale with each c-unit receiving a score for global coherence. A score of 4 indicates the c-unit is overtly related to the stimulus as defined by mention of actors/actions/objects present in the stimulus which are of significant importance to the main details of the stimulus. A score of 1 indicates the c-unit is entirely unrelated to the stimulus/topic. Subsequently, GC was estimated using the 5-point rating scaled developed by Glosser and Deser (1990) where a higher score relates to better global coherence. Finally, for each person, the semantic similarity between their stories and an exemplar story that reflected the most important story propositions was estimated using document-to-document comparison (available at http://lsa.colorado.edu/), and the average coherence (cosine) scores were calculated (for details on the mathematics for LSA see Landauer, Foltz, & Laham, 1998; Martin & Berry, 2007)).

RESULTS and DISCUSSION

Preliminary analyses have been conducted on the languages samples from 14 PWA for one of the stories (Picnic). To evaluate the concurrent validity of the tasks, Pearson correlations were conducted. All correlations among the three GC measures were strong and significant (Table 1). These findings are of great importance first because they provide evidence of construct validity**,** especially given how different the nature of LSA is from the other scales, which minimizes the probability of obtaining high correlations due to a common method factor (e.g. rating scale). Second, results are suggestive of the clinical and research potential of LSA as a tool that could be used to complement human ratings. Further, high and significant correlations were found between the GC scores and the WAB-R AQ with the exception of the Glosser and Deser scale and WAB-R AQ, $r = .44$, $p = .11$ (Table 1), but the latter finding may be due to low power.

Overall, the high correlation between the GC measures and the indices of aphasic language are indicative of a direct link between aphasia severity and GC. Further, these findings, taken in conjunction with the previous results, could be interpreted as predictive validity evidence for the GC measures. Pending analyses include: (i) exploration of divergent validity evidence for the measures for PWA, (ii) GC measurement for the language samples of NI adults and estimations of the correlations among the measures for the NI group, (iii) assessment of mean differences in GC between the groups, as measured by each tool. Clinical and theoretical implications of the results will be discussed.

REFERENCES

Armstrong, E. (2000). Aphasic discourse analysis: The story so far. *Aphasiology*, 14(9), 875-892.

Christiansen, J. A. (1995). Coherence violations and propositional usage in the narratives of fluent aphasics. *Brain and Language*, 51, 291-317.

Coelho, C., & Flewellyn, L. (2003). Longitudinal assessment of coherence in an adult with fluent aphasia: A follow-up study. *Aphasiology*, 17(3), 173-182.

Day, A. (1985). *Good dog Carl*. New York: Simon & Schuster.

Elvevåg, B. Foltz, P. W., Weinberger, D., & Goldberg, T. E. (2007). Quantifying incoherence in speech: An automated methodology and novel application to schizophrenia. *Schizophrenia Research*, 93, 304-316.

Folstein, J. A., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the mental state for the clinician. *Journal of Psychiatric Research, 12*, 189-198.

Foltz, P.W., Kintsch, W., Landauer, T.K., 1998. The measurement of textural coherence with Latent Semantic Analysis. *Discourse Processes*, 25, 285–307.

Glosser, G., & Deser, T. (1990). Patterns of discourse production among neurological patients with fluent language disorders. *Brain and Language*, 40, 67-88.

Howard, D., & Patterson, K. (1992). *Pyramids and palm trees: A test of semantic access from pictures and words.* Thames Valley: Bury St Edmunds.

Howard, M. W., Jing, B., Addis, K. M., & Kahana, M.J. (2007). In Landauer, T. K., McNamara, D. S., Dennis, S., and Kintsch, W. (Eds.), *Handbook of Latent Semantic Analysis*. (pp. 121-141). Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Kertesz, A. (2007). *Western Aphasia Battery- Revised.* New York: Grune and Stratton.

McCully, E. A. (1984). *Picnic*. Harper Collins Publishers.

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

Koutsoftas, A., Wright H. H., & Capilouto, G. J.. (2009). Discourse coherence in healthy younger and older adults. Poster presented at the Annual Meeting of the American Speech and Hearing Association, New Orleans, LA.

Landauer, T.K., Dumais, S.T., 1997. A solution to Plato's problem: the Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review* 104, 211–240.

Landauer, T.K., Foltz, P.W., Laham, D., 1998. Introduction to Latent Semantic Analysis. *Discourse Processes* 25, 259–284.

Martin, D. I., & Berry, M. W. (2007).). In Landauer, T. K., McNamara, D. S., Dennis, S., and Kintsch, W. (Eds.), *Handbook of Latent Semantic Analysis*. (pp. 121-141). Mahwah, NJ, USA: Lawrence Erlbaum Associates.

Reynolds, C. R. (2002). *Comprehensive trail making test*. Austin, TX: Pro-Ed.

Ulatowska, H. K., Olness, G., & Williams, L. J. (2004). Coherence of narratives in aphasia. *Brain and Language*, 91, 42-43.

Wechsler, D. (1997). *Wechsler memory scale - 3rd edition.* San Antonio, TX: Pyschological Corporation.

Table 1. Correlations among the GC measures and WAB-R for PWA

	GD[a]	4P[b]	LSA[c]	WAB-R AQ[d]	Spnt Score[f]
GD	1				
4P	.75**	1			
LSA	.805**	.824**	1		
WAB-R AQ	.439	.805**	.641*	1	
Spnt Score	.656**	.791**	.743**	.924**	1

[a]Glosser and Deser scale; [b]Four-point scale; [c]Latent Semantic Analysis; [d]Western Aphasia Battery – Revised Aphasia Quotient; [e] Western Aphasia Battery – Revised Spontaneous Score.

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).