Productive Vocabulary Across Discourse Types in Aphasia

Gerasimos Fergadiotis, Heather Harris Wright, & Maria Kapantzoglou
Arizona State University

Introduction

- Discourse coherence refers to the conceptual organization of discourse.
  - Global coherence (GC) is how the discourse relates to the overall topic[1,2].

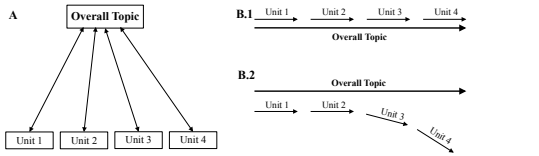


Figure 1. Conceptualizing GC. **A**. Discourse units relate to the overarching theme. The average strength of the association between each unit and the topic determines how globally coherent the discourse is. **B.1** Each discourse unit is thought of as a vector. To the extent that vectors are pointing to the same direction, they are consistent with the overall theme. **B.2** Speakers digress, or get off topic, when unit vectors point to a different direction.

- GC has been measured in people with aphasia (PWA) using different scoring methods (i.e., rating scales, coherence violations, total counts) and results have varied across studies[3,4,1,5].
- Findings may reflect
  - differences in how GC is conceptualized[6]
  - low measurement accuracy due to the lack of psychometrically sound measurement instruments. G&D was found to have lower reliability than four-p.

Purpose

1. Explore the convergent validity of (a) a new 4-Point Likert scale, and (b) a Latent Semantic Analysis (LSA) approach using the Glosser and Deser GC Scale as an external criterion.
2. Assess the instruments' predictive validity using the WAB-R AQ.

Method

Participants

	PWA (N=15)
Age (yrs)	M = 62.5 (SD=13.0)
Gender (M:F)	7:8
Education	M = 9.81 (SD=4.5)
WAB-R AQ[1]	M = 76.82 (SD=9.83)
Criteria	Chronic aphasia, L CVA, No Hx of psychiatric or neurodegenerative Dx
	Normal aided or unaided hearing and visual acuity & monolingual English speakers

[1]Western Aphasia Battery – Revised Aphasia Quotient;

Procedure

Discourse Elicitation.

- Language samples consisted of participants' storytelling of two wordless picture books.

Transcription & Language Sample Preparation.

- Samples digitally recorded & orthographically transcribed, segmented into c-units in format compatible with CLAN (MacWhinney, 2000)

Language Analyses: Global Coherence.

- Scores were estimated using the Glosser & Deser and the 4-Point scales; each C-unit received a score for global coherence and the mean was calculated by averaging across C-units and the two stories.

Glosser & Deser (1992) Global Coherence Scoring Scale

5	The utterance provides substantive information related to the general topic
4	The utterance contains multiple clauses, wherein one clause relates directly to the topic and the other relates indirectly
3	The utterance provides information possibly related to the general topic or is an evaluative statement without providing substantive information, or the topic must be inferred from the statement
2	The utterance contains multiple clauses, wherein one clause possibly relates to the general topic and one does not
1	The utterance is unrelated to the general topic or is a comment on the discourse

4-Point Global Coherence Scoring Scale

4	The utterance is overtly related to the stimulus; mention of actors/ actions/objects present in the stimulus which are of significant importance to the main details of the stimulus
3	The utterance is related to the stimulus with some inclusion of suppositional or tangential information that is relevant to the main details of the stimulus; *or* substantive information is not provided so that the topic must be inferred from the statement
2	The utterance is only remotely related to the stimulus, with possible inclusion of inappropriate egocentric information; may include tangential information or reference some element of the stimulus that is regarded as non-critical
1	The utterance is entirely unrelated to the stimulus; the utterance may be a comment on the discourse or tangential information is solely used

- The third measure is based on LSA which is a cognitive computational model of semantic memory. 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, the effects of text coherence on comprehension, and local coherence in schizophrenia.

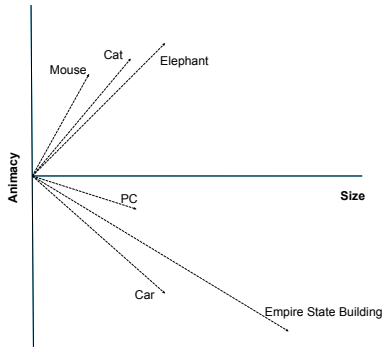


Figure 2. How does LSA work? LSA analyzes millions of words of natural discourse and decomposes their meaning into a predetermined number of latent semantic dimensions (above 2, usually 300) using Singular Value Decomposition (a matrix algebra technique related to factor analysis). Then, each unit can be represented as a point in the multidimensional semantic space. Meaning similarity can be estimated using the cosine of the unit vectors.

- LSA scores were computed using www.lsa.colorado.edu. A numerical LSA score was estimated for each C-unit based on its semantic similarity to an exemplar story; then, scores were averaged across C-units and stories...

Results

Correlations Among Measures

	G&D	Four-P	LSA
G&D	1		
Four-P	.650**	1	
LSA	.628*	.840**	1
WABR	.520*	.806**	.632*

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Regressions

Model	R	R Square	*Adjusted R Square*	F Change	$df1$	$df2$	*Sig. F Change*
G&D	0.52	0.271	0.215	4.824	1	13	0.047
Four-P	0.806	0.649	0.622	24.059	1	13	0
LSA	0.632	0.399	0.353	8.631	1	13	0.012

Discussion

- The strong correlations between the new measures and the G&D scale suggest concurrent validity evidence for the new measures.
- Correlations between the two rating scales may be somewhat inflated due to common method effects. However, the high intercorrelations of LSA with the two rating scales are not subject to common method effects and therefore are stronger evidence of concurrent validity. LSA correlated stronger with four-p than G&D.
- The best measure for predicting severity of aphasia as measured by WAB-R was Four-P (64.9%), followed by LSA (39.9%) and G&D (27.1%). Results suggest stronger evidence of predictive validity for the two new measures than G&D.
- LSA was sensitive to global coherence deviations that were also detectable by human raters.
  - This novel application of LSA provides evidence of how well the LSA algorithm can approximate human judgments based on semantic world knowledge.
- LSA can potentially be used to examine disordered language production and complement human ratings of GC.
  - Advantages of this method include objectivity, 100% reliability, and time efficiency in completing the analysis.

Limitations & Future Directions

- Rating scales are not interval scales and also yield low score variability, which may influence the results.
  - Apply models for ordered polytomous data.
- LSA does not account for word order, polysemy, anaphora, metaphor, and logical propositions.
  - Explore other models of semantic memory.
- Collect more validity evidence for different types of aphasia separately.

Acknowledgements

This work was supported by NIH/NIA Grant R01AG029476.
Contact: Gerasimos Fergadiotis – gfergadiotis@gmail.com

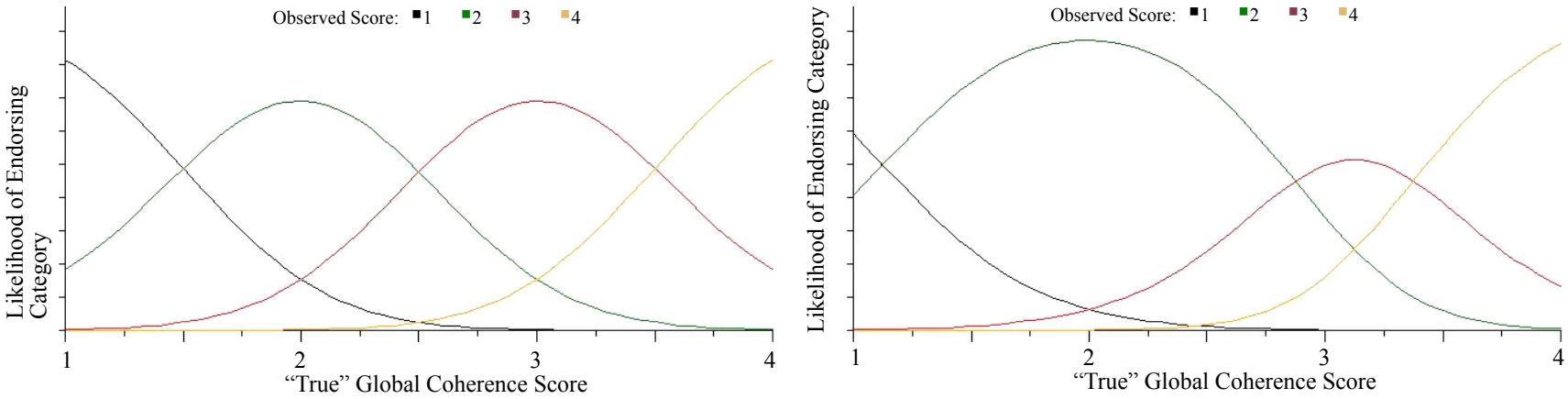
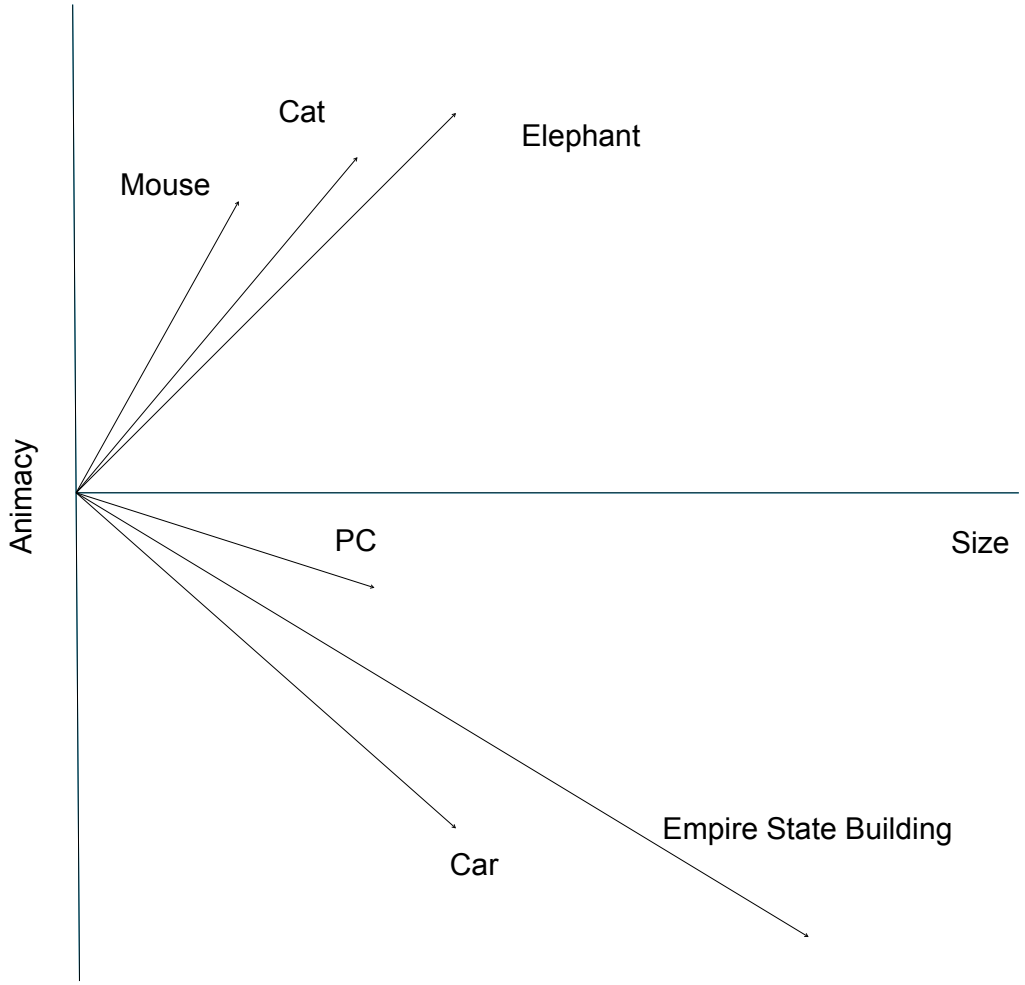


Figure 2. How do Likert Scales work? An item response theory perspective. For each discourse unit, successively higher integer scores are assigned according to explicit criteria which characterize increasing levels of global coherence. Ideally, observed scores correspond to true scores and approximate an interval scale status (left). However, often the different categories are associated with different endorsement probabilities (right) which might distort the measurement when observed scores are used.



How does LSA work?	
Step 1	"Use a training corpus typical in content and size of the people whose language is to be captured by the model."
Step 2	Create a type-by-document frequency matrix with rows for each unique word type and columns for each passage (e.g., 10^7 to 10^{10} words divided into 10^5 to 10^9 paragraphs).
Step 3	Apply a weighting function: low weight for terms that appear across many documents and high weight for terms that appear in few documents (e.g., log-entropy).
Step 4	Decompose the input matrix into orthogonal components using singular value decomposition and create a multidimensional semantic space.
Step 5	Project texts of any length to the multidimensional space and compare how similar they are using the cosine of their vectors.

Computing Story Telling Propositions

Compute propositions for story telling narrative discourse samples only

Story telling proposition - operationally defined as an event of sufficient importance to the story as a whole

Purpose - capture the extent to which participants understand and express relationships between characters, actions, and ideas.

Each wordless picture book includes a different number of propositions, as follows:

- The Picnic – 23
- Good Dog Carl - 31

Scoring: A binary scoring system is used. Responses are compared to an a priori list of main events and scored either correct (+), indicating that all the necessary information is provided or incorrect (-).

Computations:

1. Convert raw scores to proportion of story telling propositions told (SP told/total # SP)
2. Calculate proportion of story telling propositions for the following
   a. Each stimulus (Picnic, Good Dog Carl)
   b. For each session (total proportion of SP told)

Story Telling Propositions for each stimulus:

Note: Information in bold represents the essential information for each main event. Information in parentheses ( ) represents alternative ways a component of the story proposition could be stated. Information in brackets [ ] represents additional information that could have been added to complete the proposition but is not necessary for a correct (+) response.

Story Propositions may be out of order and/or 1 SP may be spread across multiple utterances in the sample: label accordingly

Good Dog Carl

1. **Mother is leaving and says, "Look after the baby Carl. I'll be back shortly."**
   - Paraphrasing acceptable if all elements of quotation are included
2. **Carl looks (**watches) **out the window & the baby climbs out of the crib** (bed) (hops on Carl's back)
   - Note final ( ): acceptable descriptions include mentioning baby climbing out of crib OR getting on Carl's back
3. **They go into the bedroom & they jump** (lay) (play) **on the bed**
   - Must include both parts of SP
4. **Get into** (go through) **the mother's makeup** (powder puff) **and jewelry**
   - Acceptable if specific terms for makeup are used (e.g. "powder puff")
5. **The dog has on a necklace** (pearls)
   - Note for SP's 4 and 5: must include more general term close to "jewelry" for SP 4 as well as specific term/description in SP 5 ("necklace")
6. **They look in the mirror**
7. **The baby is about to** (getting ready to) **go down the laundry chute so** (as) **Carl runs down the stairs.**
8. **The baby has gone through the laundry chute** (is at the bottom of the laundry chute) [waving at Carl] **Carl is at the top of the stairs** [looking at him]
   - Some mention of baby being at bottom of laundry chute/Carl being at top of stairs is required
9. **Baby gets back** (hops) **on Carl's back**
10. **They go to the living room and knock over the papers** (table) [are making a mess (destroying everything)]
    - Terms similar to "living room" or appropriate descriptions may also be acceptable such as "family room," etc.
11. **Carl's holding the back of baby's shirt while the baby is swimming in the fish tank** (aquarium)
    - Must somehow imply that Carl is holding up the baby as he swims
12. **Carl puts on music** (plays music) (turns on the stereo) **& dances (**entertains the baby) [as baby watches]
    - Must imply that Carl is dancing for the baby rather than both baby and Carl dancing
13. **The baby gets back on Carl's back** (Carl takes the baby) **into the kitchen & they open** (see what's in) **the refrigerator**
    - Note first ( ): accept other reasonable alternatives as long as it is stated/implied that Carl is taking/carrying baby to kitchen
    - Must include both parts of SP
14. **They get bread and butter & try to open the bread**
    - Mention of "bread" and "butter" in a coherent utterance sufficient for (+)
15. **They have some grapes** (the baby eats some grapes)
    - Mention of "grapes" in a coherent utterance sufficient for (+)
16. **They are thirsty so they get milk** (cream) **and Hershey's** (chocolate) **syrup**
    - Must mention both items individually: "chocolate milk" not acceptable
17. **They find** (get into) (have) **a cookie** [jar]
    - Mention of "cookie" in a coherent utterance sufficient for (+)
18. **They've made a mess** [in the kitchen (house)]
    - General statement referring to the mess or specific reference to the environment: "the baby is dirty/sticky/messy" not acceptable
19. **Carl puts the baby on his back and they go upstairs**
    - Accept other reasonable alternatives (e.g. Carl takes/brings baby upstairs) as long as it is stated/implied that Carl is taking/carrying baby upstairs
20. **Carl starts a bath** (turns on the water) **and gets soap and a washrag** (washcloth)
    - Must mention turning on water/starting bath + the 2 items
21. **He puts the baby in the tub**
22. **After the bath, Carl dries the baby off with the hair** (blow dryer)
    - Must use "blow dryer" or comparable term
23. **Carl puts the baby back in the crib**
24. **Carl cleans the house** (kitchen) (cleans up the mess)
25. **He throws away the trash & licks up the milk**
    - Note for SP's 24 and 25: must include general statement about cleaning up mess (24) AND more specific description as to how he cleans it up (25)
26. **He looks out the window to look for the mother**
27. **He makes the bed and cleans up** (straightens) **the dresser** (jewelry) (makeup)
28. [Carl sees] **the mother coming back** (up the street) (through the gate)
    - Note that Carl "seeing" the mother is optional

- Some mention of mom coming home is required

29. **Carl lays down beside the baby** (crib) (where he was) [when the mother left] (while the baby is in bed (the crib))
    - "While the baby is in bed/the crib" only is required for alternate description ("where he was")
30. **Mother says, "Good Dog Carl"** (Mother tells Carl he's a good dog)
    - No other paraphrasing acceptable
31. **She doesn't know what happened** [while she was gone]

Picnic

1. **A family of mice went for a ride** (are riding) (head off) **in their truck**
    - Mention of truck required; big, red not necessary
2. **The road is bumpy** (they hit a bump) (rock) **& the little** (baby) **mouse falls out** (off) **the back of the truck**
    - Must establish that mouse who falls out is younger/smaller
3. **No one notices and the truck continues on** (drives off)
    - Must include both parts of SP
4. **The mouse who fell out has** (is holding) **a pink stuffed animal** (bear) (mouse) **with him**
5. **The family arrives at their picnic spot** (pulls over at their destination) (the park)
6. **Mice run out of the truck** (run into the field) (park) **and** [others] **set up the picnic**
    - Must include both parts of SP
7. **The mice are playing music** (banjo)**, swinging from the trees, playing baseball.**
    - Must mention all 3 activities
8. **A couple is sitting on a blanket watching**
    - Accept reasonable alternatives (e.g. "the mom and dad/older mice are watching the others play," etc.)
9. **The little** (baby) **mouse is sad** (crying) **and alone** (by himself)**, just holding his stuffed animal** (bear) (mouse)
    - Must mention all three components of SP
10. **He notices** (sees) **some berries** (flowers)
11. [Back at the picnic] **the mouse family is playing** (having fun) **jumping** (swimming) **in the water** (lake) (pond) **& taking pictures**
    - Must mention both activities
12. **They have salad, sandwiches, & watermelon**
    - Must mention 3 or more food items; accept reasonable alternatives
13. **They are getting ready to eat** (gather around the blanket) (line up)**, when they notice** (realize) **one mouse is missing**
    - Realization of missing mouse must occur <u>after</u> statement implying that family is gathering/getting ready to eat
14. **They start searching** (looking) (calling) **for the mouse**
15. **The mice are sad** (crying) (upset)
16. **They pack up** (gather) **their belongings** (things) (picnic stuff) **& load** (go back to) **the truck**
    - Must include both parts of SP (i.e. packing up belongings AND loading/going back to truck)
17. **The little** (baby) **mouse is lying in the grass** (field) [with his stuffed animal]
18. **The mouse family drives back in the direction they came from** (retraces the road) **looking** (searching) **for the little** (baby) **mouse**
    - Must say or imply that family is traveling down <u>same</u> road from which they came

    - Must include both parts of SP (i.e. drives back/retraces + looking/searching for baby mouse)
19. **The little** (baby) **mouse is walking around in the grass**
    - Must imply that he was walking around <u>before</u> hearing family/truck
20. **The little** (baby) **mouse hears his family** (voices) **calling** (the truck) **and he runs out to the road**/voices
21. **The family and little** (baby) **mouse see each other** (reunite) **& hug** (rejoice) (embrace)
    - Must include both parts of SP (i.e. seeing each other + reaction)
22. **Little** (baby) **mouse has forgotten his stuffed animal** (something) **so he runs back** [into the field] **to get it**
23. **The family has** (decides to) **set up their picnic right there**
    - Must say or imply that picnic is taking place right there/where little mouse is found

Relationship among the 5 point Scale (Glosser & Deser, 1992), the 4 point scale and WABAQ across the stories

	5pt Picnic	5pt GDC	4pt Picnic	4pt Total
5pt GDC	.614**			
4pt Picnic	.755**			
4pt GDC		.646**	.955**	
5pt Total				.768**
4pt Total				

*$p < .01$, **$p < .001$