



Core Lexicon and Main Concept Production during Picture Description

Sarah Grace Hudspeth & Jessica D. Richardson

Neuroscience of Rehabilitation Laboratory, Department of Communication Sciences and Disorders

University of South Carolina, USA



Background

- Discourse is any number of highly individualized and complex speech acts used to transmit and receive information for survival, cooperation, and ritual purposes.^{1,2}
- Individuals with language disorders demonstrate impaired discourse ability and a resultant decline in functional communication.
- Discourse is known to be a good predictor of life participation and quality of life in persons with aphasia (PWAs).
- Analysis of discourse generally requires specialized training and can be time-consuming, limiting its clinical effectiveness.
- Researchers have suggested that analysis of a core lexicon (CoreLex) during structured narrative tasks could provide a time-efficient and informative index of functional communication abilities.³
- CoreLex has been investigated using several methods in the Cinderella story, and a procedural discourse task (how to make a peanut butter and jelly sandwich) in PWAs.
 - In two studies, CoreLex was restricted to nouns and verbs^{3,4}, while another included all parts of speech⁵.
 - For the study including all parts of speech⁵, CoreLex performance was strongly correlated to main concept production (MC; a measure of narrative adequacy) during Cinderella retelling.
- This study aimed to:
 - Determine the CoreLex of a picture sequence description task in the AphasiaBank protocol.**
 - Calculate CoreLex scores for control and PWAs.**
 - Determine how well CoreLex predicts narrative adequacy, as judged by MC analysis.**

Methods

Database:

- Transcripts of 146 control participants and 179 PWAs were retrieved from the **AphasiaBank** database.
 - 56 Anomic, 48 Broca's, 33 Conduction, 26 NABW (not aphasic by WAB), and 15 Wernicke's
- The picture description narrative was retrieved from the AphasiaBank database using the computerized language analysis (CLAN) command: **gem +sWindow +n +fWindow +d1 +t*PAR + t%mor *.cha**

Core Lexicon (CoreLex):

- The entire lexicon for **Broken Window**⁶ picture description in control participants was identified using the CLAN command: **freq +t*PAR +s"@r-*,|-*,o-%" +o *.gem.cex +d2 -s"[+exc]"**
- 993 unique lemmas were identified including all parts of speech.
- Lemmas produced by 50% or more of control participants were included in the **Broken Window** CoreLex.
 - This cutoff was selected because it yields a reasonably sized lexicon and has been used in previous research (Brown's stages of development⁷).
- CoreLex production of controls and PWAs was scored using this list.
 - Individuals received a 1 if the lemma was present in the transcript and a "0" if it was absent.
 - The sum of values across the transcript served as the CoreLex score.

Main Concepts (MC):

- Previous research identified the MCs produced by 50% of controls during **Broken Window** picture description⁸.
- Transcripts in the current study were scored using this list of 8 **Broken Window** MCs.
- A coding system (modified Kong⁸) was utilized with:
 - 0 - Absent (AB): The participant did not produce any portion of the MC.
 - 1 - Inaccurate/Incomplete (II): The participant attempted to produce a portion of the MC, but it was missing at least one essential element and another essential element was incorrect.
 - 2 - Inaccurate/Complete (IC): The participant produced a complete MC, but at least one essential element was inaccurate.
 - 2 - Accurate/Incomplete (AI): The participant produced an accurate MC, but at least one essential element was missing.
 - 3 - Accurate/Complete (AC): The participant correctly produced all essential elements.
- Scores for each MC were summed to yield the MC composite score.

Methods (continued)

Figure 1. Broken Window picture sequence elicitation stimuli⁶

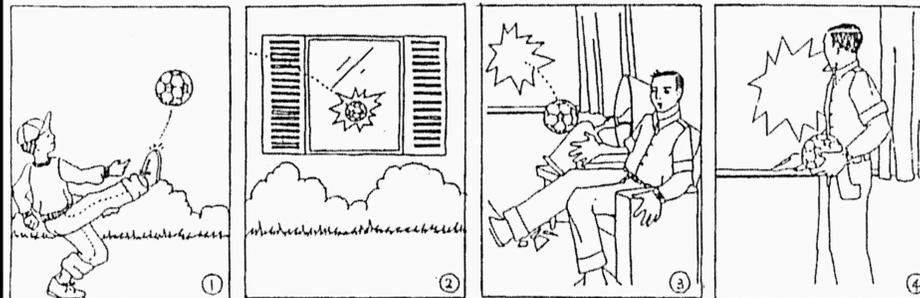


Table 1. Broken Window CoreLex list.

n boy	n out	pro:sub he	coord and
n soccer	v kick	prep of	aux be
n ball	v go	prep in	inf to
n window	v look	prep through	cop be
n lamp	pro:poss:det his	det a	
n:gerund sit	pro it	det the	

Data Analysis:

- Omnibus median tests were conducted to confirm hypothesized differences between controls and PWAs for CoreLex and MC scores.
- Planned comparisons (median tests, Holm-Bonferroni corrected) were used to identify and characterize differences between subtypes of aphasia.
- Spearman correlations were performed to investigate the relationship between CoreLex and MC scores.

Results

CoreLex:

- 22 lemmas (7 nouns, 3 verb, 3 pronouns, 3 prepositions, 2 determinants, 1 coordinator, 1 auxiliary, 1 infinitive, and 1 copula) were included in the **Broken Window** CoreLex (see Table 1).
- Omnibus median tests revealed significant differences between control participants and PWAs for CoreLex $\chi^2(5, N=325) = 195.745, p < .001$ and MCs $\chi^2(5, N=325) = 123.759, p < .001$ scores.
- Planned comparisons examining differences between aphasia subtypes were performed following recalculation of the median (excluding controls) to avoid median inflation.
- Significant differences for CoreLex and MC scores were found between the following subtypes:
 - Broca's vs. anomic – CoreLex $\chi^2 = 42.298, p < .001$ and MC $\chi^2 = 26.232, p < .001$
 - Broca's vs. conduction – CoreLex $\chi^2 = 21.811, p < .001$ and MC $\chi^2 = 10.578, p = .001$
 - Broca's vs. NABW – CoreLex $\chi^2 = 61.611, p < .001$ and MC $\chi^2 = 36.626, p < .001$
 - Broca's vs. Wernicke's – CoreLex $\chi^2 = 6.940, p = .008$ and MC $\chi^2 = 25.000, p < .001$
 - NABW vs. conduction – CoreLex $\chi^2 = 15.486, p < .001$ and MC $\chi^2 = 9.502, p = .002$
 - NABW vs. Wernicke's – CoreLex $\chi^2 = 8.464, p = .004$ and MC $\chi^2 = 7.031, p = .008$
 - NABW vs. anomic – CoreLex only $\chi^2 = 6.451, p = .011$

Table 2. Correlation of CoreLex and MC scores for each aphasia subtype.

	Omnibus	Anomic	Broca's	Conduction	NABW	Wernicke's
r_s	.818	.589	.661	.185	.474	.542
p	<.001	<.001	<.001	.295	.014	.037

Discussion

- Individuals with aphasia differed significantly from control participants on CoreLex and MC scores.
- Post hoc median tests were conducted to confirm that NABWs were significantly different from control participants for CoreLex and MC scores.
 - CoreLex $\chi^2 = 26.604, p < .001$ and MC $\chi^2 = 12.170, p < .001$
 - This result provides further evidence for the existence of a group of individuals with discourse impairments who receive little or no therapy based on standardized test scores.
- Compared to a similar investigation of the Cinderella story, CoreLex and MC scores for **Broken Window** stories more consistently differentiated fluent from non-fluent aphasia types⁵.
 - However, findings of significant differences between fluent (anomic, conduction, Wernicke's) subtypes was similar for both stories.
- For **Broken Window** and Cinderella stories only CoreLex scores significantly differed between anomic and NABW groups.
 - Although the groups may be comparable in conveying the gist of the story, they appear to differ in the typicality of the lexical items used during the telling.
- The **Broken Window** narrative is shorter than the Cinderella retelling, and may be more appropriate for some clinical settings.
- Significant correlations between CoreLex and MC scores for all subtypes except conduction aphasia indicate that a CoreLex checklist may be a time efficient and reliable predictor of narrative adequacy.
 - This may be more practicable in many instances for clinician use.
 - However, different correlation strengths for the subtypes lends support to the use of multidimensional approaches to narrative assessment.
 - MC analysis provides more detailed information about narrative adequacy, including accuracy and completeness of statements.
 - Recently developed MC checklists provide a standardized, norm-referenced and non-transcription based method of completing such a multidimensional analysis.
- It is likely that the unique deficits in conduction aphasia allow these individuals to convey much of the information about a story through circumlocution without using a typical lexicon.
 - However, this hypothesis should be confirmed through further analysis of CoreLex and MC scores in this group.

Future Directions

- Further investigations into the effect of discourse impairments on individuals who scored NABW is needed.
- The clinical relevance of CoreLex and MC measures must be established to support use as a diagnostic and therapeutic progress measure.
- In the future, practicing clinicians will be recruited to score transcripts of individuals with aphasia to ensure that results provided by these measures are valid and reliable across sites and professionals.

References

- Carey, J. W. (1988). *Communication as culture: Essays on media and society*. New York, NY: Psychology Press.
- Dimbleby, R. & Burton, G. (1998). *More than words: An introduction to communication* (3rd Ed.). New York, NY: Routledge.
- MacWhinney, B., Fromm, D., Holland, A., Forbes, M., & Wright, H. (2010). Automated analysis of the Cinderella story. *Aphasiology, 24*, 856-868.
- Fromm, D., Forbes, M., Holland, A., & MacWhinney, B. (2013). PWAs and PBJs: Language for describing a simple procedure. Poster presented at the annual Clinical Aphasiology Conference. Tucson, Arizona.
- Hudspeth, S. G., Dillow, E. P., & Richardson, J. D. (2013). Core lexicon analysis: Efficient assessment of narrative discourse in persons with aphasia. Technical research session presented at Annual Convention of the American Speech-Language-Hearing Association. Chicago, Illinois.
- Menn, L., Reilly, K.F., Hayashi, M., Kamio, A., Fujita, I., & Sasanuma, S. (1998). The interaction of preserved pragmatics and impaired syntax in Japanese and English Aphasic Speech. *Brain and Language, 61*, 183-225.
- Owens, R. E. (2008). *Language Development: An Introduction* (7th Ed.). Boston: Pearson.
- Hudspeth, S. G., Campbell, S. L., Williams, A. N., Dillow, E. P., & Richardson, J. D. (2013). Development of clinician-friendly discourse analysis tools: Main concept analysis. Poster presented at the Annual Convention of the American Speech-Language-Hearing Association. Chicago, Illinois.
- Kong, A. P. (2009). The use of main concept analysis to measure discourse production in Cantonese-speaking persons with aphasia: A preliminary report. *Journal of Communication Disorders, 42*(6), 442-464.

Acknowledgements: The authors would like to thank the Chapman Foundation for financial support, AphasiaBank developers and contributors, and Nicole Williams for her assistance in coding and scoring MCs.