Using AphasiaBank for Discourse Assessment

Davida Fromm, Ph.D.[1], **Margaret Forbes, M.A.**[2], **Audrey Holland, Ph.D.**[3], **Brian MacWhinney, Ph.D.**[1]

[1]Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania

[2]Pawley's Island, South Carolina

[3]University of Arizona, Tucson, Arizona

Abstract

AphasiaBank is a shared, multimedia database for the study of communication in aphasia. This article describes a variety of discourse measurement tools and teaching resources available at the AphasiaBank website. The discourse measurement tools include main concept analysis, core lexicon checklists, correct information unit computation techniques, and other automated analyses using the CLAN program. These tools can be used to measure a variety of aspects of language production for assessment as well as treatment evaluation and clinical research purposes. Importantly, they are intended to help make the discourse analysis process more efficient and reliable. Teaching resources include an online tutorial on aphasia, videos of typical behaviors seen in aphasia, group treatment videos, classroom activities, tutorial screencasts, and conference posters. These resources can be used for a variety of clinical and educational purposes. The AphasiaBank website is part of the larger TalkBank project which provides many other shared databases and resources that are relevant to professionals interested in communication and communication disorders.

Keywords

Aphasia; discourse; CHAT; CLAN

In 2012, this journal published an article about AphasiaBank as a resource for clinicians.[1] The article gave a brief history and description of AphasiaBank, which is a shared, multimedia database for the study of communication in aphasia. It highlighted parts of the AphasiaBank database that would be most useful to clinicians and presented a new CLAN[a] command, EVAL, designed for clinicians to take advantage of computer-based transcription and analysis tools for fast, reliable, and informative summaries of language samples. Here, we will focus on new, clinically relevant discourse measurement tools and teaching resources that have been added to the AphasiaBank website (https://aphasia.talk-bank.org/) over the ensuing years. The new measurement tools include Main Concept (MC)

Address for correspondence: Davida Fromm, Ph.D., Department of Psychology, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213 (fromm@andrew.cmu.edu).

[a]CLAN is a freely downloadable, open source software program that includes an editor for computer–based transcription and a set of programs for automated analysis of linguistic and discourse structures.[2]

checklists,[3–6] Core Lexicon Checklists,[7,8] Correct Information Unit (CIU) computation,[9] the SCRIPT command for analyzing fixed scripts,[10,11] a larger EVAL reference database, and advanced tools to compute measures from the Quantitative Production Analysis (QPA)[12,13] and Northwestern Narrative Language Analysis (NNLA)[14–16] systems. The teaching resources include an online aphasia tutorial, a set of videos highlighting examples of behaviors typically seen in the connected speech of people with aphasia (PWAs), group treatment videos, classroom activities, tutorial screencasts, and conference poster archives.

These AphasiaBank tools and resources can be particularly useful for speech-language pathologists (SLPs) who work in acute, subacute, and/or chronic treatment settings and engage in activities such as direct patient contact for assessment, treatment, reporting, and counseling; family education; staff training and supervision; interdisciplinary collaboration; clinical research projects; guest lectures; and professional outreach to the community. The focus on discourse, or connected speech, reflects its importance as the most meaningful, natural, ecologically valid, and available variety of communication. Yet, it can also be the most challenging to measure. We will present a range of measurement tools, some of which can be used without transcription and others that require transcription, to take advantage of efficient, reliable, comprehensive, and automated computer analysis programs. In addition, AphasiaBank educational resources can be used to refresh and retrain critical listening skills for lexical, syntactic, and pragmatic aspects of discourse that can inform all aspects of case management.

Bryant et al published results of a survey on clinical use of discourse analysis in aphasia assessment.[17] Although all 123 respondents said they collect spoken discourse samples for analysis, only 15.5% said they used detailed transcription-based discourse analysis for diagnosis, 16.3% for goal setting, and 14.6% for outcome measurement. Other methods used more frequently for all purposes included judgment-based discourse analysis, standardized assessment tools, and functional assessment. Factors that affected the respondents' decisions about how often to use linguistic discourse analysis included available time (the biggest factor by far), training, expertise, and resources. These results echo the findings of numerous other reports over too many years about real or perceived barriers to outcome assessment in our field.[18–21] Bryant et al recommended greater collaboration between researchers and clinicians on the topic of discourse assessment, as well as discussing how computer software can support the process of linguistic discourse analysis. The tools and resources presented here are intended to help address these factors and recommendations.

The tools described below often refer to tasks that are components of the AphasiaBank discourse protocol, which is briefly summarized in Table 1. (See article by Forbes et al[1] for more details.) All materials for elicitation of the protocol tasks are available at the AphasiaBank website in the section labeled *Protocols*; more information on all of the discourse measurement tools is available at the *Discourse Analysis* link (https:// aphasia.talkbank.org/discourse). Although some of the discourse analysis tools require transcription, which can be time-consuming, the use of automated, computer-based software makes the transcription process more efficient and reliable. For example, the CLAN editor allows for linking the media file to the transcript so that it is easy to replay individual utterances to maximize transcription accuracy. Also, CLAN transcription uses normal

English orthography (e.g., *mother's, poured, can't),* so there is no need for hand-coding of morphology (e.g., *mother/z, pour/ed, can/'t).* Once a session is transcribed, no further hand-coding (of parts-of-speech, morphology, etc.) or manual computation is needed. Those analyses can be done automatically, producing fast and accurate results. In the example below, the speaker's utterance was transcribed as part of a larger transcript. Then, one CLAN command, MOR, automatically created a morphological tier (% mor) with lexical and morphological parsing directly below each speaker utterance. The accuracy of these tags is approximately 98% for normal adult speakers.

*PAR: he kicks the ball into the neighbor's window.

%mor: pro:sub|he v|kick-3S det:art|the n|ball prep|into det:art|the adj|neighbor&dn-POSSn| window.

DISCOURSE MEASUREMENT TOOLS

Main Concept Analysis (MCA).

A main concept (MC) is an utterance with a subject, one main verb, and an object (where appropriate). According to Nicholas and Brookshire, it may contain a subordinate clause but must only have one main verb.[9] Using the large set of control participants from the AphasiaBank database, Richardson and Dalton created checklists of the essential content, or gist, for five discourse tasks in the Aphasia-Bank discourse protocol: Cinderella story (34 MCs), Sandwich task (10 MCs), Broken Window picture description (8 MCs), Refused Umbrella (10 MCs), and Cat Rescue[9] (also called "Cat in the Tree" in some literature; 10 MCs).[5,6] These checklists contain MCs mentioned by at least 33% of 92 control participants. Special symbols are used to mark a subset of the MCs in each checklist, indicating those that were spoken by 50 and 66% of the respondents. These results have been used for evaluating the performance of large samples of controls ($N = 145$) and PWAs ($N = 238$)[22] as well as a small sample ($N = 17$) of persons with primary progressive aphasia (see Dalton et al in this issue). Using a criterion of 60% or greater, Hameister and Nickels also created a list of MCs for the Cat Rescue task based on 50 control participants from the AphasiaBank database.[3] Though they used the same criterion as Richardson and Dalton,[5,6] their final list of MCs involved some additional consensus decisions made by three individual raters. Their results also provide the order in which each MC was mentioned and the exact percentage of control participants who mentioned each one. These data from the control participants were used as a benchmark for evaluating the performance of 50 PWAs on the same task, though for this analysis the authors made some minor procedural changes (e.g., accepting both recognizable verbal and nonverbal responses, not utilizing the scoring system).

Without labor-intensive transcription of discourse samples, clinicians can use any of these five discourse tasks and MC checklists to get an objective measure of a PWA's ability to provide "essential content," as defined by responses to the same tasks from a normative sample. Scoring could be done live or, preferably, from a recording or simple orthographic transcription. In addition to an overall score for number of MCs mentioned, more detailed qualitative scoring can be done for accuracy and completeness yielding a composite score, as described in the articles. Dalton and Richardson showed that these MC checklists were

sensitive to differences between PWAs and controls, and reported on other MCA research demonstrating reliability in the scoring and test–retest properties.[22] Hameister and Nickels reported that the PWAs produced fewer MCs than the controls and had greater variability in the order of MCs.[3] The MC checklists and information about MC scoring are available at AphasiaBank's *Discourse Analysis* link.

Core Lexicon.

Core lexicons are a way to capture how many of the most typical words used by a normative group for a particular task are used by a comparison group. In fact, this issue features several articles on the use of core lexicons and their application to aphasia, extending the scope of earlier work that examined the top 10 nouns and verbs in participants' responses to the procedural discourse and Cinderella story telling tasks from the AphasiaBank discourse protocol (see articles by Dalton et al and Kim et al in this issue).[23–27] Dalton and Richardson analyzed responses to the Broken Window picture description task from the AphasiaBank discourse protocol.[7] They created a core lexicon (CoreLex) list of the 24 unique lemmas (word roots, such as "run" for "runs," "running" and "ran") used by at least 50% of the normative sample (92 control participants). When transcripts from 235 PWAs were analyzed to see how many of those 24 CoreLex words were used in their picture descriptions, results revealed significant differences between PWAs and controls as well as between all aphasia subtypes. CoreLex lists are also available for the Cat Rescue and Refused Umbrella tasks,[7] and will soon be available for the Cinderella story. Again, these materials can be found at the *Discourse Analysis* link on the AphasiaBank website.

CIU Computation.

The aforementioned article that examined SLPs' use of discourse analysis for assessing PWAs reported that CIU analysis was by far the most frequently used (24.5% of respondents) of any analysis method defined in the research literature.[17] CIU analysis measures the communicative informativeness and efficiency of connected speech based on a set of scoring rules.[9] It has long been shown to have high test–retest reliability and to adequately reflect listeners' perceptions of informative discourse in structured discourse tasks.[28,29] CLAN software now includes an automated CIU command that computes number and proportion of CIUs as well as CIUs per minute. To do those computations, the language sample needs to be (1) transcribed in a CHAT file (or converted to a CHAT file from a text file or SALT transcript); (2) coded with "[e]" to exclude certain words from the CIU word count that cannot be excluded automatically (see the Appendix of Nicholas and Brookshire[9]); and (3) linked to the audio or video file if computation of efficiency measures (words/minute, CIUs/minute) is desired. While most of the CIU exclusion rules have been automated, a few still require hand coding. Specific rules about what is automatically excluded from total word and total CIU counts can be found in the CIU section of the *Discourse Analysis* webpage. Words followed by [e] will be ignored in the CIU computations as per the rules. If a string of words is to be excluded, that string must be enclosed in angle brackets and then followed by [e]. If a single word is to be excluded, angle brackets are not needed. Excluded words will NOT be ignored for the total word count, only the CIU word count. Here are some examples of coded lines of a Cookie Theft picture description in CHAT format with CIU exclusion markings.

1. Exclusion of vague or nonspecific word:

   *PAR: mother is drying one of those things [e].

2. Exclusion of qualifier for unambiguous information:

   *PAR: <I think that> [e] boy give girl something [e].

3. Exclusion of incorrect information:

   *PAR: the girl [e] is on the stool.

4. Exclusion of commentary on the task:

   *PAR: <this is hard> [e].

CIU analysis with structured discourse tasks can be a valuable tool for aphasia assessment purposes as well as measurement of change over time. Automation of the majority of CIU computation is currently being implemented, leaving only a minimal amount of manual exclusion coding that will be necessary to add to the transcript. Once tested and completed, the details will be available at the AphasiaBank website's *Discourse Analysis* page and in the CLAN manual.

SCRIPT.

The SCRIPT command was developed for analysis of fixed scripts used in assessment, treatment, or research. It has been used to evaluate the results of script training programs,[10,11] and is currently being used to evaluate oral reading test performance. The analysis compares the participant's performance to that of the model (therapy script, reading passage, etc.) and provides the following information in a spreadsheet: number and percent of correct words, number and percent of omitted words, number of added words, number of recognizable errors, number of unrecognizable errors, number of utterances with unintelligible content, and number of missing utterances. Recognizable errors such as "brella" for "umbrella" can be transcribed with target replacements, which in CHAT format would be brella [: umbrella]. Unrecognizable errors are words that did not have any obvious relation to the expected target word or would not be recognizable as such to an average listener. Those are transcribed with an error code next to them, indicating that the word was a neologism with an unknown target [* n:uk] or a semantic (real-word) error with an unknown target [* s:uk]. Unintelligible content is transcribed as xxx. The SCRIPT command also outputs a list of the omitted words, added words, and errors in a text file format. A transcription shortcut for this type of analysis with less severely impaired individuals is simply to use the model script transcript as a template and then modify it as needed for each participant or each time it is used. In other words, duplicate the model script, play the audio or video for the participant, and modify each line based on what the participant produced. This can be a useful way to compile data over time on a given speaker. It can also be an efficient way to compile group data on many individuals' performance on a reading passage or other type of fixed script. Further information on the SCRIPT command can be found in the Analysis Commands section of the CLAN manual (https://talkbank.org/manuals/CLAN.pdf).

EVAL.

The EVAL command, described in detail in a previous issue of this journal by Forbes et al,[1] is an example of a CLAN command that creates a composite profile, in this case, with 34 outcome measures relevant to aphasia (e.g., mean length of utterance [MLU] in words, type–token ratio, words per minute, open/closed class word ratio, number or percentage of word errors, propositional idea density, number of verbs per utterance). The EVAL command has been used in numerous publications, theses, and conference presentations to analyze aspects of discourse in a variety of populations, such as aphasia, primary progressive aphasia, traumatic brain injury, and second language acquisition.[30–38] The original publication about EVAL showed how to compare an individual's pretherapy and posttherapy performance on a discourse task. However, another application of this EVAL command is to compare an individual's performance on a particular discourse task (e.g., Cinderella story) to the performance of others with the same aphasia type or to individuals without aphasia. As can be seen in the CLAN dialog box in Fig. 1, the comparison options also include age ranges and gender. So, if an individual with a Western Aphasia Battery-Revised Aphasia Quotient (WAB-R AQ) of 93.0 and a WAB-R subtype of anomic aphasia is compared with the entire AphasiaBank database of PWAs with that type of aphasia, the results appear as illustrated in the Excel spreadsheet in Fig. 2. The output shows that the reference database includes 134 PWAs with anomic aphasia who did the Cinderella task. The asterisks in the spreadsheet indicate where the individual is 1 (*) or 2 (**) standard deviations above or below the mean of the reference database. For example, he was 2 standard deviations above the mean database value for total number of utterances (77 compared with 31.672) and 1 standard deviation above the mean database value for total number of word errors (26 compared with 9.007). The results for this variable as well as parts of speech can be computed as raw numbers (as was done here) or as percentages of total words (which is the command default). The SLP's guide (https://talkbank.org/manuals/Clin-CLAN.pdf) and the tutorial screencasts (discussed later) provide step-by-step instructions on how to create transcripts and run the EVAL command. Support is also available from AphasiaBank for individual or small group training using computer screensharing options.

C-QPA and C-NNLA.

CLAN also includes single commands for more advanced automated grammatical analyses using the QPA and several sections of the NNLA systems.[9,14–16] The C-QPA command produces an analysis worksheet with approximately 17 outcome measures for each sentence in the transcript and a summary worksheet with data from the sentence-by-sentence analysis worksheet plus the computations derived from all those numbers. The C-NNLA command currently computes approximately 50 measures, including general language measures (e.g., MLU, number of utterances, and number of words), lexical variables (e.g., number and percentage of nouns, verbs, and adjectives), morphological variables (e.g., number of regular and irregular past tense and plural forms), utterance level variables, and sentence-level variables. However, these analyses require more extensive CHAT transcriptions (e.g., with full error coding) and more extensive individualized evaluation than do any of the other discourse measurement tools discussed here. Because of this, they are less likely to be utilized by busy clinicians. Readers interested in learning more about these commands and

analyses can find information at the Aphasia-Bank webpage (*Discourse Analysis* link) and in the CLAN manuals.

TEACHING TOOLS AND RESOURCES: APHASIA, CLAN

Grand Rounds and Examples.[b]

These two links at the AphasiaBank webpage contain material that emphasizes discourse characteristics from a wide variety of individuals with aphasia. *Grand Rounds* is an online, guided tutorial on aphasia, focusing on how language differs across aphasia types and language tasks. Types of aphasia covered in the *Grand Rounds* include anomic, Broca, conduction, global, transcortical motor, transcortical sensory, and Wernicke. Tasks highlighted in the video segments include stroke stories, free speech, naming tasks, Cinderella story narratives, picture descriptions, repetition tasks, and procedural discourse. The *Grand Rounds* text has short case histories and discussion questions built around 40 video segments from dozens of PWAs. Additional "Treatment Focus" questions stimulate thinking and discussions about ways to approach intervention to improve communication for each individual case presentation. In addition to its obvious applicability for graduate classes on aphasia, this resource could be used in the clinic to refresh one's own skills in aphasia assessment and treatment, to facilitate orientation and training with volunteers or newer SLPs, and to help increase awareness about aphasia for colleagues in related health care professions. Current AphasiaBank members also use the aphasia videos with aphasia support groups.

The *Examples* link provides very short video clips of common features from the connected speech of PWA at the word level, sentence level, and discourse level. The page is organized by features (e.g., phonemic paraphasia, circumlocution, agrammatism, empty speech), with a description and several links to video examples of the feature. Each video link includes captions and basic information about the individual (e.g., WAB-R AQ, aphasia type). Again this resource can be used in classrooms and other educational settings as well as within clinical settings for review and training purposes.

Classroom Activities.

This link downloads a Word file with ideas and exercises for clinical assessment and treatment planning, measuring different aspects of discourse (e.g., CIUs, MLU), using the EVAL command described earlier, coding speech errors (e.g., phonemic paraphasias, semantic paraphasias) and other behaviors (e.g., MCs), and comparing across aphasia types as well as across other disorders (e.g., right hemisphere disorder). Several of these exercises have been contributed by AphasiaBank members, and we appreciate and encourage this type of resource sharing. Once again, these activities can be used for individual or group study and practice purposes.

---

[b]These resources are password protected. Licensed SLPs, educators, and researchers who would like access can send an email request to Brian MacWhinney (macw@cmu.edu) with contact information, affiliation, and a brief general statement about how they envision using the resources.

Group Treatment Videos.

A large collection of aphasia group treatment videos from six different sites in North America is available from a link at the *Database* section of the website. Clicking on the *Group Treatment* link will open a page with an index to the videos organized by contributors. From there, clicking on the corpus name (e.g., BU) will open a page with more information about the contributors and the videos. From any corpus page, clicking on the Browsable Database link will open a page with simple instructions on how to see and hear videos. A directory of the videos will appear in the top left corner of the page in the Browsable Database. To watch, just click on the video of interest (e.g., 6_25_Starbucks.cha in the BU corpus) and then press the play arrow on the video that appears in the lower left corner. The video can be controlled from that player and can be made full screen by clicking the farthest right icon (with four diagonally oriented arrows) next to the volume control. These group treatment videos could be used for a variety of educational and research purposes in classrooms and clinics. Though these videos are not transcribed, they provide a range of material on group treatment approaches for PWA of different abilities. They also demonstrate the enjoyment that participants have engaging in this type of activity. This is an area ripe for more attention in clinical education and research.

Tutorial Screencasts.

Over 40 screen-casts—https://talkbank.org/screencasts/—provide short tutorials on how to do many different CLAN functions including using the Browsable Database to watch videos, starting a CHAT transcript from scratch, time-linking a transcript to the media file, converting a text file to a CHAT transcript, and running a variety of CLAN commands (e.g., EVAL). These screencasts can be used by clinicians and students at any level (e.g., undergraduate, graduate) and in any area of study in the field (e.g., fluency, child language) to facilitate the use of reliable and efficient automated linguistic analysis tools for discourse. They were designed as an adjunct to the more in-depth, written manuals available at the main TalkBank webpage (https://talkbank.org/).

Posters.

A final resource to highlight is a collection of 69 (as of August 2019) posters that were based on AphasiaBank data and presented at a variety of conferences. Clicking on the *Posters* link at the AphasiaBank website opens a page listing the poster titles and authors by conference and year. Clicking on the link in the *Authors* column opens the actual poster that was presented. These materials provide information that is otherwise unavailable for people who did not attend those conferences or missed seeing those particular posters. Additionally, the posters can stimulate ideas for research questions to be asked and investigated in classroom and clinic settings.

In summary, the information provided here highlights the newest materials that are likely to be most relevant and useful for clinicians, clinical researchers, and clinical educators who may not yet be very familiar with all these resources. It is by no means a comprehensive review of all the materials available at the AphasiaBank website. Interested readers are encouraged to browse the website and request membership (as explained earlier), if they want to request access to password protected links. AphasiaBank is part of the larger

TalkBank project (https://talkbank.org/) that includes similar data collections of relevance to SLPs in FluencyBank, RHDBank, TBIBank, ASDBank, DementiaBank, and PhonBank. Readers are encouraged to explore all of these resources available from the TalkBank Project.

FINANCIAL DISCLOSURE

REFERENCES

1. Forbes MM, Fromm D, MacWhinney B. Aphasia-Bank: a resource for clinicians. Semin Speech Lang 2012;33(03):217–222 [PubMed: 22851343]

2. MacWhinney B The Childes Project: Tools for Analyzing Talk: Vol. II: The Database. Mahwah2000

3. Hameister I, Nickels L. The cat in the tree–using picture descriptions to inform our understanding of conceptualisation in aphasia. Lang Cogn Neurosci 2018;33(10):1296–1314

4. Nicholas LE, Brookshire RH. Presence, completeness, and accuracy of main concepts in the connected speech of non-brain-damaged adults and adults with aphasia. J Speech Hear Res 1995;38(01):145–156 [PubMed: 7731205]

5. Richardson JD, Dalton SG. Main concepts for three different discourse tasks in a large non-clinical sample. Aphasiology 2016;30(01):45–73

6. Richardson JD, Dalton SG. Main concepts for two picture description tasks: an addition to Richardson and Dalton, 2016. Aphasiology 2019:1–8. DOI: 10.1080/02687038.2018.1561417 [PubMed: 33012945]

7. Dalton SG, Richardson JD. Core-lexicon and main-concept production during picture-sequence description in adults without brain damage and adults with aphasia. Am J Speech Lang Pathol 2015;24(04):S923–S938 [PubMed: 26140462]

8. Tanaka T, Branch W, Dalton SG, Richardson JD. It's raining cats: further development of discourse checklists for picture description tasks. Poster presented at the American-Speech-Language-Hearing Association Convention; November 2016; Philadelphia, PA

9. Nicholas LE, Brookshire RH. A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. J Speech Hear Res 1993;36(02):338–350 [PubMed: 8487525]

10. Fridriksson J, Hubbard HI, Hudspeth SG, et al. Speech entrainment enables patients with Broca's aphasia to produce fluent speech. Brain 2012;135 (Pt 12):3815–3829 [PubMed: 23250889]

11. Szabo GB, Fromm D, Heimlich T, Holland A. Script training and its application to everyday life observed in an aphasia center. Paper presented at the Clinical Aphasiology Conference; May 2014; St. Simons Island, GA

12. Rochon E, Saffran EM, Berndt RS, Schwartz MF. Quantitative analysis of aphasic sentence production: further development and new data. Brain Lang 2000;72(03):193–218 [PubMed: 10764517]

13. Saffran EM, Berndt RS, Schwartz MF. The quantitative analysis of agrammatic production: procedure and data. Brain Lang 1989;37(03):440–479 [PubMed: 2804622]

14. Thompson CK, Shapiro LP, Tait ME, Jacobs B, Schneider S, Ballard K. A system for the linguistic analysis of agrammatic language production. Brain Lang 1995;51(01):124–129

15. Thompson CK, Cho S, Hsu CJ, et al. Dissociations between fluency and agrammatism in primary progressive aphasia. Aphasiology 2012;26(01):20–43 [PubMed: 22199417]

16. Thompson CK. Northwestern narrative language analysis (NNLA) theory and methodology. Available at: http://anr.northwestern.edu/wp-content/uploads/2016/09/NNLA_REVISED.Fall2015.pdf. Accessed October 14, 2019

17. Bryant L, Spencer E, Ferguson A. Clinical use of linguistic discourse analysis for the assessment of language in aphasia. Aphasiology 2017;31(10):1105–1126

18. Armstrong E Aphasic discourse analysis: the story so far. Aphasiology 2000;14(09):875–892

19. Bryant L, Ferguson A, Spencer E. Linguistic analysis of discourse in aphasia: a review of the literature. Clin Linguist Phon 2016;30(07):489–518 [PubMed: 27002416]

20. Duncan EA, Murray J. The barriers and facilitators to routine outcome measurement by allied health professionals in practice: a systematic review. BMC Health Serv Res 2012;12(01):96 [PubMed: 22506982]

21. Simmons-Mackie N, Threats TT, Kagan A. Outcome assessment in aphasia: a survey. J Commun Disord 2005;38(01):1–27 [PubMed: 15475012]

22. Dalton SGH, Richardson JD. A large-scale comparison of Main Concept production between persons with aphasia and persons without brain injury. Am J Speech Lang Pathol 2019;28(1S):293–320 [PubMed: 31072179]

23. MacWhinney B, Fromm D, Holland A, Forbes M, Wright H. Automated analysis of the Cinderella story. Aphasiology 2010;24(6-8):856–868 [PubMed: 25067870]

24. Fromm D, Forbes M, Holland A, MacWhinney B. PWAs and PBJs: language for describing a simple procedure. Poster presented at the Clinical Aphasiology Conference; May 2013; Tucson, AZ

25. Kim H, Wright HH. Tutorial on core lexicon: development, use, and application. Semin Speech Lang 2020;41:20–31 [PubMed: 31869846]

26. Dalton SGH, Hubbard HI, Richardson J. Moving toward non-transcription based discourse analysis in stable and progressive aphasia. Semin Speech Lang 2020;41:32–44 [PubMed: 31869847]

27. Dalton SGH, Kim H, Richardson J, Wright HH. A compendium of core lexicon checklists. Semin Speech Lang 2020;41:45–60 [PubMed: 31869848]

28. Pritchard M, Hilari K, Cocks N, Dipper L. Reviewing the quality of discourse information measures in aphasia. Int J Lang Commun Disord 2017;52(06):689–732 [PubMed: 28560767]

29. Doyle PJ, Tsironas D, Goda AJ, Kalinyak M. The relationship between objective measures and listeners' judgments of the communicative informativeness of the connected discourse of adults with aphasia. Am J Speech Lang Pathol 1996;5(03):53–60

30. Boyle M Stability of word-retrieval errors with the AphasiaBank stimuli. Am J Speech Lang Pathol 2015;24(04):S953–S960 [PubMed: 26151900]

31. Casilio M, Rising K, Beeson PM, Bunton K, Wilson SM. Auditory-perceptual rating of connected speech in aphasia. Am J Speech Lang Pathol 2019;28(02):550–568 [PubMed: 31136232]

32. Colantuoni DE. The Effects of Transcranial Direct Current Stimulation on Narrative Abilities in Primary Progressive Aphasia. Master's Thesis. College Park, MD: University of Maryland; 2018

33. Faroqi-Shah Y, Milman L, Pranger M, Damele D, Leiva A, Neuharth M. Identifying reliable discourse measures for second language speakers of English. Presentation at the annual convention of the American Speech-Language-Hearing Association, Philadelphia, PA; November 2016

34. Fromm D, Greenhouse J, Hou K, et al. Automated proposition density analysis for discourse in aphasia. J Speech Lang Hear Res 2016;59(05):1123–1132 [PubMed: 27657850]

35. Hsu CJ, Thompson CK. Manual versus automated narrative analysis of agrammatic production patterns: the northwestern narrative language analysis and computerized language analysis. J Speech Lang Hear Res 2018;61(02):373–385 [PubMed: 29411008]

36. Loughnane M, Murray L. Communicative efficiency and executive function in normally aging and older adults with a mild acquired brain injury. Aphasiology 2018;32(01):130–132

37. Stark BC. A comparison of three discourse elicitation methods in aphasia and age-matched adults: implications for language assessment and outcome. Am J Speech Lang Pathol 2019;28(03):1067–1083 [PubMed: 31181170]

38. Thorne J, Faroqi-Shah Y. Verb production in aphasia: testing the division of labor between syntax and semantics. Semin Speech Lang 2016;37(01):023–033

Learning Outcomes:

As a result of this activity, the reader will be able to (1) measure discourse in aphasia more efficiently and reliably using a variety of discourse analysis tools presented here; (2) compare the linguistic discourse performance of any given person with aphasia to that of a large reference database of people with and without aphasia; and (3) use AphasiaBank's educational resources for training and counseling purposes in classrooms and clinical settings.

Figure 1.
EVAL dialog box.

	File_DB	Duration_(sec)	Total_Utts	MLU_Words	FREQ_types	FREQ_tokens	FREQ_TTR	Words_Min	Verbs_Utt
1	File_DB	Duration_(sec)	Total_Utts	MLU_Words	FREQ_types	FREQ_tokens	FREQ_TTR	Words_Min	Verbs_Utt
2	thompson01a.cha	889	77	6.805	151	533	0.283	35.973	1.234
3	+/-SD	00:00:03.281	2.024	-0.017	1.613	1.889	-1.249	-0.631	0.291
4		••	••		•	•	•		
5	Mean Database	00:04:21	31.672	6.839	79.896	220.306	0.427	54.014	1.107
6	Min Database	00:00:24	2	2.5	4	6	0.178	8.182	0.12
7	Max Database	00:20:37	127	12.875	209	801	0.842	146.878	2.234
8	SD Database	00:03:11.319	22.394	2.019	44.069	165.505	0.115	28.608	0.434
9									
10		Word_Errors	Utt_Errors	density	Nouns	Plurals	Verbs	Aux	3S
11	thompson01a.cha	26	29	0.402	106	23	103	25	16
12	+/-SD	1.539	2.222	-0.903	2.458	2.822	1.752	3.794	0.678
13		•	••		••	••	•	••	
14	Mean Database	9.007	9.03	0.456	37.44	6.478	42.351	5.373	8.806
15	Min Database	1	1	0.28	1	3	2	5	3
16	Max Database	52	51	0.664	152	35	167	30	53
17	SD Database	11.044	8.989	0.061	27.897	5.855	34.62	5.173	10.614
18									
19		1S_3S	PAST	PASTP	PRESP	prep	adj	adv	conj
20	thompson01a.cha	24	54	4	21	43	20	34	2
21	+/-SD	4.289	2.366	0.934	4.335	2.507	1.599	1.492	-0.405
22		••	••		••	••	•	•	
23	Mean Database	3.739	15.731	1.97	3.522	12.873	7.56	14.612	3.836
24	Min Database	5	2	1	3	1	1	1	4
25	Max Database	25	78	11	23	49	37	55	24
26	SD Database	4.724	16.178	2.173	4.031	12.017	7.779	12.99	4.536
27									
28		det	pro	noun_verb	open_closed	#open-class	#closed-class	retracing	repetition
29	thompson01a.cha	83	62	1.116	0.746	223	299	78	94
30	+/-SD	3.398	1.007	-0.127	0.198	2.255	1.716	3.983	3.341
31		••	•			••	•	••	••
32	Mean Database	21.567	33.306	1.217	0.696	83.478	129.597	14.47	17.672
33	Min Database	1	1	0.278	0.384	3	3	1	1
34	Max Database	108	119	6	2.833	320	463	109	143
35	SD Database	18.076	28.499	0.79	0.251	61.865	98.715	15.951	22.845
36									
37	+/- SD * = 1 SD, ** = 2 SD								
38	Database Gems:Cinderella								
39	Database keywords:Anomic								
40	#_files_in_database:134								

Figure 2.
EVAL spreadsheet output.

Table 1

AphasiaBank Standard Discourse Protocol[*]

Free speech	Stroke story and coping Important event
Picture description	Broken window Refused umbrella Cat rescue (also called "cat in the tree")[9]
Story narrative	Cinderella
Procedural	Peanut butter and jelly sandwich

[*] Protocol script and stimuli are available at: https://aphasia.talkbank.org/protocol/.