Discourse measurement in aphasia research: have we reached the tipping point?

Aimee Dietz & Mary Boyle

Routledge
Taylor & Francis Group

Check for updates

FORUM TARGET ARTICLE

Discourse measurement in aphasia research: have we reached the tipping point?

Aimee Dietz[a] and Mary Boyle[b]

[a]Department of Communication Sciences & Disorders, University of Cincinnati, Cincinnati, OH, USA;
[b]Department of Communication Sciences & Disorders, Montclair State University, Bloomfield, NJ, USA

ABSTRACT

Background: Over the past few decades, clinical aphasiologists have increasingly included various measures of discourse, across a variety of genres, to assess treatment effects and to understand spoken language in people with aphasia. The paradigm shift from assessing language at the word and sentence level to assessing at the discourse level shows an increased awareness of capturing meaningful outcomes; however, this has led to a proliferation of new discourse metrics. These new metrics often lack reliability, validity, and stability data to support their use. In fact, we propose that a tipping point has been reached.

Aims: To review the recent move toward establishing a basic core outcome set (COS) for aphasia. In support of this effort, we make a strong argument for the need to dovetail these efforts with the development of a core outcome set for discourse (D-COS).

Main contribution: We make a case that the current state of the art in aphasia-related discourse research leaves researchers unsure of the most representative outcomes to include in future investigations and clinicians with little guidance on best practice. We applaud and support the recent efforts toward the development of a general COS for aphasia; however, we propose that the need for a D-COS is urgent because the increasing numbers of discourse measures make it difficult to replicate and compare outcomes across studies. Moreover, these issues likely interfere with adaptation of the use of discourse measurement in clinical practice.

Conclusions: Many health-related fields have adopted a COS to guide their research and ensure that the field's data can be compiled, compared, and contrasted for a better overall understanding of how treatments address targeted illnesses. We argue that, in addition to the recent progress toward a COS for aphasia, a D-COS is also required to improve our understanding of how our treatments improve the discourse of people with aphasia, which in turn will aid in evoking a paradigm shift in clinical practice. To be successful, it is essential that the development of a D-COS includes a process by which all stakeholders are included: people with aphasia, caregivers, clinicians, and clinical researchers.

CONTACT Aimee Dietz ✉ dietzae@ucmail.uc.edu 📧 Department of Communication Sciences & Disorders, University of Cincinnati, PO Box 670379, 3202 Eden Avenue, Cincinnati, OH 45267-0379, USA

The evolution of discourse measurement in aphasia

Duffy (2000) asserted that paradigms influence the way we think about aphasia, about our research, and about our clinical work. Early-on, speech-language pathologists typically focused assessment at the word and sentence levels (Bottenberg, Lemme, & Hedberg, 1985; Ulatowska, Macaluso-Haynes, & North, 1980). Then, researchers began to investigate how aphasia affects spoken language production across a variety of genres, from procedural to narrative retells and to conversational discourse. The early studies focused largely on aspects of the macrolinguistic structure of discourse, like coherence, cohesion, and story grammar (e.g., Whitworth et al., 2015). Following this, aspects of microlinguistic structure, like lexical and syntactic production, were examined (e.g., Pashek & Tompkins, 2002; Saffran, Berndt, & Schwartz, 1989). As we learned more about how aphasia affects discourse, investigators began to assess whether treatment conducted at the single-word or single-sentence level (e.g., Boyle & Coelho, 1995; Schwartz, Saffran, Fink, Myers, & Martin, 1994), discourse level (e.g., Goral & Kempler, 2009; Greenwood, Grassly, Hickin, & Best, 2010), or at multiple levels in the same treatment program (e.g., word, sentence, and connected speech) (Milman, Vega-Mendoza, & Clendenen, 2014; Whitworth et al., 2015) improved discourse production.

The paradigm shift from a focus on single words and sentences to the inclusion of a variety of discourse genres reflected the fact that discourse is the currency of most adults' everyday communication. This new paradigm influenced, at least in part, an attempt to develop more ecologically valid discourse outcome measures for aphasia treatment. Despite the paradigm shift toward the use of discourse, rather than word or sentence-level tasks, no standardized, widely accepted set of discourse outcome measures exists. As such, researchers rely on existing measures, or develop new ones, often without validity, reliability, or stability data to support their use (Boyle, 2014, 2015). This has resulted in a proliferation in the number and type of discourse outcome measures. Just as one example, a cursory review of the aphasiology literature revealed four measures of informativeness: (1) correct information units (Nicholas & Brookshire, 1993), (2) information units (McNeil, Doyle, Fossett, Park, & Goda, 2001), (3) lexical information units (Andreetta & Marini, 2015), and (4) units of new information (Del Toro et al., 2008), each with its own slightly different operational definition. Do we need to have four different ways to measure informativeness in discourse? Undoubtedly, each group of investigators who developed these metrics did so because they felt that their particular operational definition of informativeness was important for their particular study. However, since each of these metrics is slightly different, it makes comparison across the studies difficult. This creates a quandary for the next group of investigators who want to include a measure of informativeness in their study: which of the four should they choose, or should they develop a new measure uniquely suited to the aims of their own investigation?

We recognize that investigators need to be able to develop measures to answer their research questions if no suitable measures exist. However, it seems that we have reached a tipping point. A tipping point occurs when a threshold is crossed, and an idea, trend, or behavior spreads like wildfire (Duffy, 2000; Gladwell, 2000). Clearly, the idea of assessing discourse in aphasia has crossed a threshold. The question is whether the proliferation of new, often idiosyncratic metrics lacking reliability, validity, and

stability data to support their use is the best way to improve our understanding of assessment and treatment of discourse in aphasia.

We are excited to see a paradigm shift toward the use of discourse as the gold standard for assessing spoken language production; however, we have found that it is not without challenges. For example, as clinical researchers, we aim to select outcome measures that will be sensitive to changes in spoken discourse. But with so many to choose from, we struggle to determine the most promising measures that are also valid and stable. Since relatively few discourse outcome measures provide information about their stability and validity, should we limit our choices to those that do? Furthermore, how do we respond to the amazement of our neurology colleagues when we tell them that there is no core outcome set (COS) for aphasia research that they can use to evaluate the outcome measures in the grant proposals they are asked to review? In order to strengthen the influence of our research in the wider research community, we must come to a consensus on the key outcomes that represent meaningful change in discourse production. As academic faculty, we also struggle with how to evoke this same paradigm shift in clinical practice. It is paramount that the "Ivory Tower" analogies be put to bed and that the work we so passionately execute can translate into clinically relevant results, which can be easily implemented by clinicians. Transcription and coding issues aside, how do we encourage clinicians and students alike to incorporate discourse into their assessment repertoire when there is such an overwhelming number of outcome choices that we (the "experts") are struggling to make sense of ourselves?

The solution: development of a COS

The paradigm shift of assessing spoken language at the discourse level shows an increased awareness by aphasiologists regarding the need to capture meaningful outcomes; however, to address the aforementioned tipping point dilemma, we propose the careful development of a discourse COS (D-COS). According to the Core Outcome Measures in Effectiveness Trials (COMET) a freely available, online COS database (http://www.comet-initiative.org/) (COMET, 2016) a COS can serve as a mechanism to increase the evidence base at a collective level (Brady et al., 2014; Wallace, Worrall, Rose, & Le Dorze, 2014a, 2014b). In essence, a COS includes the minimum standard set of outcome measures required to capture treatment-related change, without restricting scholars from developing outcomes pertinent to their study question. Evidence from other health field areas, such as asthma (National Improvement Partnership Network (NIPN), 2016), arthritis (Kirkham, Boers, Tugwell, Clarke, & Williamson, 2013), and medical management of stroke (National Institute of Neurological Disorders and Stroke (NINDS), 2016) (to name a few), document how the use of a COS has improved the ability to compare and combine research outcomes and facilitate increased transparency in reporting of treatment studies (Kirkham, Gargon, Clarke, & Williamson, 2013; Sautenet, Caille, Halimi, Goupille, & Giraudeau, 2013). All of this would, undoubtedly, be equally useful for research in clinical aphasiology. Currently, 50 fields are participating in COMET, including aphasiology. However, as of this writing, the search term "aphasia" yielded only four published studies representing the preliminary work toward development of a general COS for aphasia (Wallace et al., 2014a, 2014b; Wallace, Worrall, Rose, & Le Dorze, 2016; Wallace et al., 2016).

In a recent Discussion Forum in *Aphasiology*, Wallace and her colleagues argued the need for a COS in aphasia research (Wallace et al., 2014a) and explained the function of their *Improving Research Outcome Measurement in Aphasia* project. To date, this project has garnered the support of the international aphasiology community, gathered input from stakeholders (Wallace et al., 2016, 2016), and initiated a systematic review. This work is timely given that the most recent Cochrane review of aphasia treatment studies reported an overwhelming 44 different outcomes (including measures for discourse) for 57 investigations (Brady, Kelly, Godwin, Enderby, & Campbell, 2016)! The wide range of reported outcomes is due to the large variety of domains targeted by investigators, including, but not limited to (1) functional communication, (2) comprehension: auditory, reading, and gestural; (3) expression: spoken (repetition to generative discourse) and written (copying text to generative discourse); (4) life participation, (5) quality of life; and (6) aphasia severity. Developing a COS for aphasia treatment research is a tremendous undertaking and, given the broad spectrum of reported outcomes, we applaud Wallace and her colleagues for their leadership. In fact, it is very likely that subsets of COS will be required to properly address all of these domains; thus, our proposal to develop a D-COS. We discussed and debated this notion with attendees at a recent Clinical Aphasiology Conference Round Table (Boyle & Dietz, 2015). The Round Table participants conveyed frustration regarding the arduous task of selecting the most appropriate spoken discourse measures to reflect change as a result of treatment. As such, they unanimously agreed on the importance of forging ahead with the development of a D-COS, in support of the work by Wallace and colleagues to establish a general COS for aphasia treatment.

Our aim is to engage clinical researchers who focus on discourse in aphasia to begin thinking about, talking about, and acting on the dire need to develop a D-COS for spoken discourse. We urge conversation and debate regarding the need to establish subsets of a D-COS to address the various levels of discourse structure (e.g., microstructure, macrostructure) across a wide variety of spoken discourse genres (e.g., narrative, procedural, conversational, etc.). We also think that it is critical that careful consideration be given to how written discourse fits into the overall picture, especially given the surge in social media and other text-based forms of communication in the 21st century. We are not naïve about the difficulty of this enterprise and the obstacles that must be overcome. We realize that the clinical research community comprises only one set of stakeholders in this enterprise. However, we believe that if we can agree on a basic D-COS (or subcategories of a D-COS for different discourse structures and genres) that can be used across studies, we can then engage in discussions with other stakeholders to determine the overlap in outcome preferences and adjust the D-COS to represent all stakeholder needs. In closing, we cannot emphasize strongly enough the importance of developing a D-COS. In the long run, being able to accumulate, compare, and contrast data across studies can only improve our understanding of how our treatments improve the discourse of people with aphasia. This, in turn, will aid in evoking a paradigm shift in clinical practice toward the use of discourse measures to examine changes in spoken language. Ultimately, the goal is for all stakeholders to work toward an agreed-upon D-COS that generates discourse-related meaningful outcomes and improved life participation for people with aphasia.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Andreetta, S., & Marini, A. (2015). The effect of lexical deficits on narrative disturbances in fluent aphasia. *Aphasiology*, *29*, 705–723. doi:10.1080/02687038.2014.979394

Bottenberg, D., Lemme, M. L., & Hedberg, N. (1985). Analysis of oral narratives of normal and aphasic adults. In R. H. Brookshire (Ed.), *Clinical Aphasiology Conference Proceedings, 1985*, pp.241–247. Minneapolis, MN: BRK Publishers. Retrieved from http://aphasiology.pitt.edu/archive/00000860/01/15-31.pdf.

Boyle, M. (2014). Test-retest stability of word retrieval in aphasic discourse. *Journal of Speech, Language, and Hearing Research*, *57*, 966–978. doi: 10.1044/2014_JSLHR-L-13-0171

Boyle, M. (2015). Stability of word-retrieval errors with the AphasiaBank stimuli. *American Journal of Speech-Language Pathology*, *24*, S953–S960. doi: 10.1044/2015_AJSLP-14-0152

Boyle, M, & Coelho, C. A. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology*, *4*(4), 94-98. doi: 10.1044/1058-0360.0404.94

Boyle, M., & Dietz, A. (2015). *Discourse measurement in aphasia research: Have we reaching the tipping point?* Paper presented at the Clinical Aphasiology Conference, Monterey, California.

Brady, M. C., Ali, M., Fyndanis, C., Kambanaros, M., Grohmann, K. K., Laska, A., … Varlokosta, S. (2014). Time for a step change? Improving the efficiency, relevance, reliability, validity and transparency of aphasia rehabilitation research through core outcome measures, a common data set and improved reporting criteria. *Aphasiology*, *28*, 1385–1392. doi:10.1080/02687038.2014.930261

Brady, M. C., Kelly, H., Godwin, J., Enderby, P., & Campbell, P. (2016). Speech and language therapy for aphasia following stroke (Article CD000425). Cochrane Database of Systematic Reviews 2016. Retrieved August 1, 2016, from John Wiley & Sons, Ltd. 10.1002/14651858.CD000425.pub4/full

COMET. (2016). Core outcome measures in effectiveness trials initiative (COMET). Retrieved from http://www.comet-initiative.org/

Del Toro, C. M., Altmann, L. J. P., Raymer, A. M., Leon, S., Blonder, L. X., & Rothi, L. J. R. (2008). Changes in aphasic discourse after contrasting treatments for anomia. *Aphasiology*, *22*, 881–892. doi:10/1080/02687030701844204

Duffy, J. R. (2000) Paradigms and tipping points. Paper presented to the 2000 Clinical Aphasiology Conference, Kona Coast, HI.

Gladwell, M. (2000). *The tipping point*. New York: Little, Brown and Company.

Goral, M., & Kempler, D. (2009). Training verb production in communicative context: Evidence from a person with chronic non-fluent aphasia. *Aphasiology*, *23*, 1383–1397. doi:10.1080/02687030802235203

Greenwood, A., Grassly, J., Hickin, J., & Best, W. (2010). Phonological and orthographic cueing therapy: A case of generalized improvement. *Aphasiology*, *24*, 991–1016. doi:10.1080/02687030903168220

Kirkham, J. J., Boers, M., Tugwell, P., Clarke, M., & Williamson, P. R. (2013). Outcome measures in rheumatoid arthritis randomised trials over the last 50 years. *Trials*, *14*, 1–8. doi:10.1186/1745-6215-14-324

Kirkham, J. J., Gargon, E., Clarke, M., & Williamson, P. R. (2013). Can a core outcome set improve the quality of systematic reviews?–A survey of the co-ordinating editors of cochrane review groups. *Trials*, *14*, 1–5. doi:10.1186/1745-6215-14-21

McNeil, M. R., Doyle, P. J., Fossett, T. R. D., Park, G. H., & Goda, A. J. (2001). Reliability and concurrent validity of the information unit scoring metric for the story retelling procedure. *Aphasiology*, *15*, 991–1006. doi:10.1080/02687040143000348

Milman, L., Vega-Mendoza, M., & Clendenen, D. (2014). Integrated training for aphasia: An application of part-whole learning to treat lexical retrieval, sentence production, and discourse-level communications in three cases of nonfluent aphasia. *American Journal of Speech-Language Pathology*, *23*, 105–119. doi:10.1044/2014AJSLP-12-0054

National Improvement Partnership Network (NIPN) (2016). Asthma core measures: Core and optional process and outcome. Retrieved from https://www.uvm.edu/medicine/nipn/documents/NIPNAsthmaMeasures.pdf

National Institute of Neurological Disorders and Stroke (NINDS) (2016). NINDS common data elements. Retrieved from https://commondataelements.ninds.nih.gov/#page=Default

Nicholas, L. E, & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech Language and Hearing Research*, *36*, 338-350. doi: 10.1044/jshr.3602.338

Pashek, G. V., & Tompkins, C. A. (2002). Context and word class influences in lexical retrieval in aphasia. *Aphasiology*, *16*, 261–286. doi:10.1016/j.sbspro.2013.09.028

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, *37*, 440–479. doi: 10.1016/0093-934X(89)90030-8

Sautenet, B., Caille, A., Halimi, J. M., Goupille, P., & Giraudeau, B. (2013). Better reporting and greater homogeneity in outcome measures are seen in randomized trial protocols when guidelines exist. *Journal of Clinical Epidemiology*, *66*, 838–846. doi:10.1016/j.jclinepi.2013.03.015

Schwartz, M. F., Saffran, E. M., Fink, R. B., Myers, J. L., & Martin, N. (1994). Mapping therapy: A treatment programme for agrammatism. *Aphasiology*, *8*, 19–54. doi: 10.1080/02687039408248639

Ulatowska, H. K., Macaluso-Haynes, S., & North, A. J. (1980). Production of narrative and procedural discourse in aphasia. In R. H. Brookshire (Ed.), *Clinical Aphasiology Conference Proceedings, 1980*, pp.17–27. Minneapolis, MN: BRK Publishers. Retrieved from http://aphasiology.pitt.edu/archive/00000684/01/10-02.pdf

Wallace, S. J., Worrall, L., Rose, T., & Le Dorze, G. (2014a). A good outcome for aphasia. *Aphasiology*, *28*, 1400–1404. doi:10.1080/02687038.2014.935119

Wallace, S. J., Worrall, L., Rose, T., & Le Dorze, G. (2014b). Measuring outcomes in aphasia research: A review of current practice and an agenda for standardisation. *Aphasiology*, *28*, 1364–1384. doi:10.1080/02687038.2014.930262

Wallace, S. J., Worrall, L., Rose, T., & Le Dorze, G. (2016). Which treatment outcomes are most important to aphasia clinicians and managers? An international e-Delphi consensus study. *Aphasiology*, 1464–5041. doi:10.1080/02687038.2016.1186265

Wallace, S. J., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., … Gauvreau, C. A. (2016). Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disability and Rehabilitation*, 1–16. doi:10.1080/09638288.2016.1194899

Whitworth, A., Leitão, S., Cartwright, J., Webster, J., Hankey, G. J., Zach, J., … Wolz, V. (2015). NARNIA: A new twist to an old tale. A pilot RCT to evaluate a multilevel approach to improving discourse in aphasia. *Aphasiology*, *29*, 1345–1382. doi:10.1080/02687038.2015.1081143