How to Analyze Verbal Protocols

Thora Tenbrink (t.tenbrink@bangor.ac.uk)
School of Linguistics and English Language
Room 331, 3rd floor New Arts,
Bangor University, College Road, Bangor, Gwynedd
LL57 2DG, UK
Fon: +44 1248 382263; Fax: +44 1248 383267

Keywords: Think-aloud protocols; verbal data; cognitive processes; cognitive discourse analysis; linguistic structure; problem solving; complex cognition.

Objectives and Scope

Cognitive science researchers are interested in a subject that is not directly accessible to observation: processes in the mind and brain, thoughts and thought processes. One way of addressing higher-level cognitive processes is to analyze verbal protocols produced along with cognitively complex tasks (Ericsson & Simon, 1993), such as problem solving or decision making. Linguistic data of this kind can be seen as an external representation of some aspects of what is going on in the mind. In particular, think-aloud protocols and retrospective reports provide procedural information that complements other data, such as decision outcomes and behavioral performance results.

This tutorial explores the scope and limitations of verbal protocol analysis, and offers practical support for systematic analysis procedures. Language data can be analyzed with respect to *content* as well as *structure*. Conventionally, the focus of verbal protocol analysis lies on the content of verbal data, addressing those aspects (e.g., particular thought processes or strategies) that the speakers are themselves aware of (or 'heed', Ericsson & Simon, 1993). The content-based inspection of verbal reports, particularly if carried out by experts in the problem domain and set against a substantial theoretical background (Krippendorff, 2004), often leads to well-founded specific hypotheses about the cognitive processes involved.

A detailed linguistic analysis can substantially support such content-based insights, but it can also offer further insights (e.g., Hölscher et al., 2011; Tenbrink et al., 2011; Tenbrink & Seifert, 2011; Tenbrink & Wiener, 2009). Research in cognitive linguistics, psychology, discourse analysis, and psycholinguistics indicates that patterns in language are systematically related to patterns of thought (e.g., Chafe, 1998). Drawing on these insights, one focus of the tutorial is to identify types of linguistic *structure* that point to specific cognitive processes. This is the main idea in the method of *Cognitive Discourse Analysis* (CODA) (Tenbrink, 2008; Tenbrink & Gralla, 2009; Tenbrink, 2010).

Some aspects of language use reflect cognitive aspects that go beyond conscious reflection by individual speakers, and that are not necessarily directly observable in linguistic content. Speakers are typically unaware of the cognitive structures that are reflected in particular ways of framing a representation linguistically. Furthermore, they are not consciously aware of the *network of options* (Tenbrink & Freksa, 2009) that allows for a range of linguistic choices beside their own, which emerges more clearly by considering a larger data set collected under controlled circumstances. According to previous research in cognitive linguistics and discourse analysis (e.g., van Dijk, 2008), linguistic features such as the verbal representation of semantic domains reflected in ideational networks, specific choices of prepositions, lexical omissions and elaboration, conceptual perspectives revealed by language, presuppositions, hesitation and discourse markers, and many other linguistic features indicate certain conceptual circumstances; these are related to the current cognitive representations in ways that distinguish them from other options available in the network. In particular, the chosen linguistic options reflect what speakers perceive as sufficiently relevant to be verbalized, as well as the information status assigned to the diverse parts of the verbalization.

Besides building on established insights about the significance of particular linguistic choices, validating evidence for the relationship between patterns of language use and the associated cognitive processes can be gained by triangulation, i.e., the combination of linguistic analysis with other types of evidence such as memory or behavioral performance data, reaction times, eye movements, decision outcomes, or any other relevant data that can be collected in cognitively complex tasks.

Format and organization

This tutorial is designed to cover a half day (three hours) and will be highly interactive. The tutorial will take the participants' current or intended projects as a starting point. It will be organized so as to cover the complete process of language data analysis (from initial ideas to evaluation of analysis results), including short presentations, discussion, and practical exercises where feasible. In particular, the following issues will be addressed:

Motivation: How (and to what extent) can language data serve as empirical resources to address research questions in cognitive science?

Data collection: What kinds of issues need to be considered in the light of actual research purposes?

CODA based analysis (main part): Systematic data annotation and interpretation informed by linguistic insights.

Triangulation: How can other types of empirical data complement the insights gained from language?

Participants who have already collected natural language data are encouraged to bring examples as handouts or on their computers. Furthermore they are encouraged to contribute a 10-min talk related to one step of this process, and also to raise questions or issues to discuss for other steps. It is envisioned to prepare either a collection of papers or a collective paper, with authors interactively developing content based on combinations of their talks and the discussed issues.

Target audience information

There is no prerequisite for taking this tutorial. It is open for researchers in cognitive science at any point in their career, ranging from graduate students to established experts.

Participants interested in a future publication are encouraged to submit a 300-word abstract to propose a 10-minute presentation as part of the tutorial, and / or a critical issue to discuss.

Tutor Information

Thora Tenbrink is a Lecturer in Cognitive Linguistics at Bangor University (Wales, UK), and a principal investigator in two projects in the Collaborative Research Center SFB/TR 8 Spatial Cognition (Bremen/Freiburg, Germany). Her main interest concerns complex cognitive processes and their representation in language. She is the author of "Space, Time, and the Use of Language" (Mouton de Gruyter, 2007), and co-editor of "Spatial Language and Dialogue" (Oxford University Press, 2009) and "Representing space in cognition: Interrelations of behavior, language, and formal models" (Oxford University Press, in press). Current research addresses cognitive strategies in various problem solving tasks, spatial communication in complex built environments, cognitive transformation processes, and inferences derived by problem solvers from situational clues, experience, and verbal and graphical information. See http://knirb.net for further information.

Previous instantiations

This tutorial has previously been offered in various versions as listed below (see Tenbrink et al., 2012, for a report). The current version will focus on complex problem solving processes across all areas of cognitive science, tailored to the needs of its participants by establishing email contact in advance as far as possible.

"Understanding spatial thought through language use". Half-day tutorial at *Spatial Cognition,* August 31 - September 03, 2012, Abbey Kloster Seeon, Germany.

"Understanding cognitive processes through language use". Half-day tutorial at *ICCM 11th International Conference on Cognitive Modeling*, April 12-15, 2012, Berlin, Germany.

Workshop "Language analysis in cognitive science". Cognitive Science Institute, University of Osnabrück (Germany), May 7-8, 2011.

Course "Language analysis in cognitive science" at the Cognitive Science Institute, University of Freiburg (Germany), summer semester 2009.

References

Chafe, W. (1998). Language and the Flow of Thought. In M. Tomasello (ed.), *The New Psychology of Language. Cognitive and Functional Approaches to Language Structure*. Mahwah: Lawrence Erlbaum.

van Dijk, T.A. (2008). *Discourse and Context. A Sociocognitive Approach*. Cambridge: Cambridge University Press.

Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.

Hölscher, C., Tenbrink, T., & Wiener, J. (2011). Would you follow your own route description? *Cognition 121,* 228-247.

Krippendorff, K. (2004). *Content Analysis: An Introduction to its Methodology* (2nd ed.). London: Sage.

Tenbrink, T. (2008). The verbalization of cognitive processes: Thinking-aloud data and retrospective reports. In W. Ramm & C. Fabricius-Hansen (eds.), *Linearisation and Segmentation in Discourse. Multidisciplinary Approaches to Discourse (MAD 08)*, Feb 20-23 2008, Lysebu, Oslo. Oslo: Dept. of Literature, Area Studies and Europ. Languages, Univ. of Oslo.

Tenbrink, T. (2010). CODA: Kognitive Diskursanalyse. In: E. Ruigendijk, T. Stolz, & J. Trabant (Hrsg.), *Linguistik im Nordwesten: Beiträge zum 1. Nordwestdeutschen Linguistischen Kolloquium*. Bochum: Brockmeyer.

Tenbrink, T., Bergmann, E., & Konieczny, L. (2011). Wayfinding and description strategies in an unfamiliar complex building. In L. Carlson, C. Hölscher, & T.F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Tenbrink, T., D'Odorico, T., Hertzberg, C., Mazman, G., Meneghetti, C., Reshoeft, N., & Yang, J. (2012). Tutorial report: Understanding spatial thought through language use. *Journal of Spatial Information Science 5*:107-114. http://josis.org/index.php/josis/article/viewFile/124/90

Tenbrink, T., & Freksa, C. (2009). Contrast sets in spatial and temporal language. *Cognitive Processing 10 Supplement 2*, S322-S324.

Tenbrink, T., & Gralla, L. (2009). Accessing complex cognitive processes via linguistic protocol analysis. In U. Schmid, M. Ragni, M. Knauff (Eds.): *Proceedings of the KI 2009 Workshop Complex Cognition*, Paderborn, Germany, September 15, 2009.

Tenbrink, T., & Seifert, I. (2011). Conceptual Layers and Strategies in Tour Planning. *Cognitive Processing 12:1*, 109–125.

Tenbrink, T., & Wiener, J. (2009). The verbalization of multiple strategies in a variant of the traveling salesperson problem. *Cognitive Processing 10*:2, 143-161.