Book review

Melissa Bowles (2010). *The Think-Aloud Controversy in Second-Language Research.* New York & London: Routledge | ISBN 978-0-41599-484-2

Reviewed by: *Arnt Lykke Jakobsen, Copenhagen Business School, Denmark*

Previous volumes in Routledge's powerful Second Language Acquisition Research Series include such classics as Gass & Mackey's *Stimulated Recall Methodology in Second Language Research* (2000), the same authors' *Data Elicitation for Second and Foreign Language Research* (2007), and McDonough & Trofimovitch's *Using Priming Methods in Second Language Research* (2008), all of them very useful methodological guide books. Melissa Bowles' volume continues this fine tradition.

The volume has four short and two long chapters (two and three). This, together with the title of the volume, suggests that the book was conceived around an interest in the controversy over the use of think aloud (TA) (Ch. 2) and in arguments used for vindicating the use of TA (Ch. 3). However, the interest of the book goes beyond this.

The first chapter surveys the history of the think aloud method with particular focus on its use in L1 and L2 research, where the method has also been used as a means of eliciting information about thought processes and cognitive processing involved in the performance of language tasks. There is also a brief look at some areas outside second language acquisition (SLA), including accounting, economics and market research.

The analysis undertaken in chapters 2 and 3 is followed by very practical advice in chapter 4, e.g. on how to administer think aloud experiments, how participants should be instructed, what types of language experiments are best suited for think aloud data collection, what language(s) (L1 and/or L2) should be used to verbalize thought, and on how validity is best assured. Similarly, in chapter 5, the author offers valuable advice on a number of data analysis considerations, including how to transcribe verbal data at different levels of granularity, and how to ensure that data are representative of whatever is being investigated. And she goes on to illustrate different ways in which data can be coded, depending on what analytical categories are brought to bear on the material.

journal of WRITING RESEARCH

Jakobsen, A.L. (2012). The Think-Aloud Controversy in Second-Language Research [Book Review]. *Journal of Writing Research, 4*(1), 107-110. http://dx.doi.org/10.17239/jowr-2012.04.01.5
Contact and copyright: Earli | Arnt Lykke Jakobsen, Copenhagen Business School Copenhagen, Dalgas Have 15, 2Ø.073, DK - 2000 Frederiksberg, Denmark| alj.ibc@cbs.dk.

The very short sixth chapter returns to the think aloud controversy theme. The central controversy is diagnosed as existing between cognitivists and Vygotskyan learning theorists. Researchers with an interest in cognition and information processing tend to think of concurrent think aloud data as a window into cognitive processes. Their ambition is to demonstrate that think aloud can produce relevant data for exploring cognition *without* noticeable 'reactivity'. Their concern is that think aloud might be reactive, i.e. might change the primary cognitive processes they wish to study, which would invalidate the method in their eyes. In Vygotsky's view of the function of human language, verbalization – especially in the form of inner speech ('egocentric speech') – serves to help regulate cognitive processes and generate new knowledge. The opportunity to verbalize what we are engaged in mediates the internalization of knowledge. In other words, learning emerges through verbalization, and therefore educationalists see think aloud as *necessarily reactive*.

Chapter 2 first gives a detailed report on ten studies concerning the reactivity of non-metacognitive verbal reports that were not discussed by Ericsson and Simon in their introduction to the second edition of *Protocol Analysis* (1993). Nine of the studies treated non-verbal tasks or word-level tasks, such as anagrams. Only one studied the effect of non-metacognitive TA on a continuous language task (Stratman and Hamp-Lyons' study of an L1 revision task (1994)). The results of these studies are quite mixed: Six studies found reactivity for latency, four found no such effect. This is in perfect agreement with the theory as propounded by Ericsson and Simon, who predicted that concurrent non-metacognitive TA might have a latency effect due to queuing of information in short-term memory. Only one study found reactivity for accuracy (negative for two groups of participants, but positive for one group). In the reviewer's study of the effect of non-metacognitive TA on the performance of a translation task (2003), it was found that in addition to delaying production (cf. also Krings, 1995; 2001), TA also degraded the quality of processing.

Chapter 2 continues with a report on 30 studies, all from the field of cognitive psychology, investigating the reactivity of metacognitive TA – again on the performance of non-verbal tasks. Here, also, findings were quite mixed. The majority of studies found reactivity for both accuracy and latency. With respect to latency, whether measured as time-on-task or as reaction time to individual task items, nearly all studies found that metacognitive TA delayed performance. With respect to accuracy, the picture was again mixed: in some cases metacognitive TA improved performance, and in some cases it was found to hinder performance.

A number of questions arise from the survey in chapter 2, most of which are addressed in chapter 3, the main question being to what extent results obtained in experiments with non-verbal tasks can be predicted to apply in experiments with (continuous) verbal tasks, in L1 or L2 or both, such as reading, writing and translation.

Chapter 3 opens with a report on 11 recent studies of L2 learners engaged in L2 tasks. The results of these studies are also quite mixed. In the majority of studies there was an effect on latency with TA, but not in all studies. With respect to accuracy or

effect on performance, some studies found that concurrent TA improved task performance, while others found that it hindered performance. In the face of such apparent contradictions, Bowles proposes to undertake a meta-analysis of the studies aimed at establishing the validity of TA on verbal (L2) tasks by answering the following main research question: "Are think-alouds reactive for accuracy and/or latency when used in conjunction with verbal tasks?" (p. 78). The methodological procedure followed is carefully explained. We are informed that a meta-analysis "compares outcomes of a range of studies with an array of independent variables, in an attempt to identify patterns" (p. 78). The array of independent variables includes such factors as type of verbal report (non-metacognitive vs. metacognitive), language of verbal report (L1, L2 or both), language of task (L1 vs. L2), type of task (reading vs. writing vs. grammar learning vs. meta-language), and L2 proficiency (beginning vs. intermediate vs. advanced). This array of variables suggests that that there may not be a simple answer to the main research question, but that it will have to be qualified by a specification of how much reactivity and what kind of reactivity (positive or negative) can be predicted to occur, given a certain configuration of the above variables.

After careful coding of the 11 studies (plus 3 more), and after testing for group homogeneity, Bowles is able to calculate effect sizes using Cohen's *d*. For each independent variable the effect size is calculated for (1) reading, (2) receptive form (i.e. grammar) learning, (3) productive form learning, and (4) latency. Very few results are reported for writing. In the majority of cases, Bowles is unable to find statistically reliable results, and overall the effect sizes found are weak. The only statistically reliable and predictable effect that seems to manifest itself across all independent variables is the effect of metacognitive TA on latency. Other effects that achieve statistical significance by the criteria used are linked to a specific independent variable. If the type of task is reading, then concurrent TA can be predicted to slow down the process. If the type of task is writing, then think aloud can be predicted to hinder production, etc.

Though such findings are only very partial answers to the main research question, they are highly valuable both from a research perspective and from the point of view of teaching. If we know to what extent co-variance can be demonstrated between such factors as L1 or L2 proficiency, task type, task language, task design, participants' professional expertise level on the one hand and positive or negative reactivity with non-metacognitive or metacognitive TA on the other, we can both design experiments more confidently and apply such knowledge pedagogically to enhance learning.

The Ericsson/Simon vs. Vygotsky conflict remains unresolved at the end and more research is clearly needed to establish the validity of TA on verbal tasks more certainly and in greater detail, but Bowles' book succeeds well in preparing the ground for much more focused research in the field.

References

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

Jakobsen, A. L. (2003). Effects of Think Aloud on Translation Speed, Revision and Segmentation. In: Alves, F. (ed.) *Triangulating Translation. Perspectives in Process Oriented Research*. Amsterdam/Philadelphia: John Benjamins, pp. 69-95. doi: 10.1075/btl.45.08jak

Krings, H. P. (1995). *Texte reparieren. Empirische Untersuchungen zum Prozeß der Nachredaktion von Maschinenübersetzungen.* Habilitationsschrift, Universität Hildesheim, Hildesheim.

Krings, H. P. (2001). *Repairing Texts: Empirical Investigations of Machine Translation Post-Editing Processes* (G. S. Koby, Trans.). Kent: Kent State University Press.

Stratman, J. F., & Hamp-Lyons, L. (1994). Reactivity in concurrent think-aloud protocols: Issues for research. In P. Smagorinsky (Ed.), *Speaking about writing: Reflections on research methodology* (Vol. 8, pp. 89-111). Thousand Oaks, CA: Sage.