

DIARY STUDIES IN CLASSROOM SLA RESEARCH PROBLEMS AND PROSPECTS

John Fry

Abstract

This short article is an appraisal of *Diary Studies* (DS) as a research method. DS have a contribution to make to research into *second language acquisition* (SLA)¹ but advocates of DS occasionally omit discussion of problems in deriving generalisable conclusions from them. In what follows therefore the author identifies three claims made for DS and then goes on to discuss, in relation to these claims, the problems of (a) gathering valid data; (b) interpreting the data; and (c) deciding whether or not DS are a type of *non-interventional* research. He ends by explaining two ways in which DS can be beneficial despite their shortcomings and one way (action research as distinct from mainstream research) in which the shortcomings become positive advantages.

Claims for Diary Studies

Various claims have been made for *diary studies* (DS). These are:

- Claim 1: that they serve to generate hypotheses about classroom *second language acquisition* (SLA),
- Claim 2: that they provide insights into learner variables (i.e. learning strategies, and affective and psychological factors involved in learning),
- Claim 3: that they can provide insights into the processes of SLA itself.

In terms of illuminating the nature of SLA, these claims become successively stronger (claim 3, of course, being the

strongest). Also the stronger claims appear to assume the weaker ones. It is clear that Matsumoto (1987) embraces all three:

Each of the diary studies . . . contains unique and noteworthy information which contributes to our understanding of the processes underlying second language learning and teaching in a formal classroom setting. (p. 21)

Problems of Gathering Data

Terminology

Following Schumann (1977) and Bailey (1980), Matsumoto (1987) distinguishes the terms *introspective* and *non-introspective* data. For them the term *introspective* is the term given to data where the researcher and the diarist are the same person; *non-introspective* where the researcher and diarist(s) are different people. For others, including Cohen and Hosenfeld (1981), Mann (1982) and Cohen (1983), *introspective* data is that which is gathered from subjects while they carry out a task; an example is the *think aloud* protocols used by Cohen (1983) in gathering data on reading strategies. *Retrospective* data is that which is collected after the event. I shall be using this latter terminology. DS will be considered as examples of retrospective data.

Introspective Data

Introspective data holds a far stronger claim than retrospective data to be capable of providing valid insights into SLA, as it represents no additional load on long-term memory. Approaches to introspective data presuppose that learners have a high degree of what Mann (1982) refers to as *metacognitive awareness*, that is, they are able to accurately "observe the contents of their minds and infer from this process in operation" (p. 89). In considering during-task verbal reports, Seliger (1983) argues that although awareness of processes might be available for a few seconds, it will be obscured by the task.

Reporting on how one is doing a task while doing it is a double task. This obviously increases the cognitive load and is likely also to affect the processes. Furthermore, learners can only report what they are conscious of, the unconscious processes remaining inaccessible. How can we be sure that when a learner reports using strategy X it wholly represents what is going on? Moreover, the nature of the task might itself suggest the use of certain strategies, so that what is reported of processes may in fact be inferred from the task, amounting perhaps to what Seliger (1983) calls "post hoc guessing." What is most frustrating is that we can not assess the scale of these problems from the data. We can not be sure just how problematic they are. Nevertheless they remain as concerns that must affect the status of the data and what it may be taken as evidence of.

Seliger (1983) concludes of introspective verbal reports that they can be useful generating hypotheses about learning (claim 1) and with regard to learner variables (claim 2), but nothing direct can be expected regarding learning processes (claim 3). Cohen (1983), though, admits that such data primarily reflects attitudes to and strategies for learning (claim 2) rather than conscious learning processes.

Retrospective Data

With retrospective data, all the problems of cognition in introspective data are magnified by the lapse of time between the event and the reporting of it. The longer this time is the greater the load on memory capacity, and therefore the greater the likelihood of a report being a piecemeal and edited account of the event. Perceptions change with time and people are apt to forget. By its very nature, such data involves levels of unconscious editing. What may have seemed salient during the event is quite likely to differ from what is recorded as salient later. The event will be summarised. It is therefore difficult to see how DS can provide insights into the processes of SLA. Claim 3 must therefore be untenable.

Researcher Intervention

One argument advanced in favour of DS is that the diarist can be free from research intervention, free to record his/her perceptions without prompting, and that, further, the value of DS lies in this freedom from constraints. However, even if the researcher tries to eschew all intervention, waiting until the diary is complete before looking at it, there are still problems that threaten the validity of the data.

Consistency, in terms of time (i.e. that the data is recorded at a fixed time after the event, preferably as soon as possible) and in terms of depth (i.e. the level of detail recorded) cannot be ensured. Diary-keeping is time-consuming and initial enthusiasm may give way to fatigue. And lack of consistency will diminish the potential usefulness of the data.

Even where researchers intend to limit the dangers of subjects' editing of data by remaining inexplicit about the goals of their research, there is still the danger that subjects will try to guess these intentions and provide what they think the researcher is after or simply try to show themselves in a good light. And, of course, editing is not necessarily conscious.

Another problem is that the act of recording aspects of learning behaviour will raise consciousness of that behaviour and may change it. For example, comparison of early and later diary entries may reveal changes in the learners' perceptions of, say, reactions to peer correction or use of a particular strategy for vocabulary acquisition. While this is of interest as regards claims 1 and 2 (hypotheses about learning and learner variables), it can yield nothing direct as regards learning processes (claim 3). This consciousness-raising aspect of diary-keeping will be considered further in the later discussion of different research uses of diary data.

If the above problems exist when the researcher adopts a stance of non-intervention, any degree of overt intervention, such as periodic examining and discussion of a diary with a subject, will increase them.

They do not render claims 1 and 2 empty, but they do indicate that DS data cannot be expected to yield any firm statements. Rather they will produce tentative hypotheses (claim 1) and interesting possibilities (claim 2).

Cohen (1983) argues that introspective data on SLA processes may be obtainable if subjects are trained (i.e. in metacognitive awareness) to provide the required kinds of data. The problem here is that the training is highly likely to affect both task performance and the processes themselves. Mann (1982) echoes this concern. In discussing the use of practice tasks to refine subjects' reporting, she warns that:

. . . subject training may bias the data towards the experimenter's desires and expectations, whereas no training may result in the loss of potentially relevant information through the subject's ignorance of interesting features, limited metacognitive awareness or through the subject's discomfort and unease with the experimental task.
(p. 91)

Problems of Interpretation

Perhaps the first question regarding the analysis of the data is, as Mann (1982) says:

Do we approach the data with pre-defined categories or do we allow the data to drive the analysis? (p. 95)

Clearly, if we are interested in the learners' perceptions of their learning, we would be wise to adopt the latter course. Putting the data through a sieve of pre-established categories risks forcing a particular interpretation onto the data and overlooking points of possible interest. But how specific should the categories that emerge be? The best course is probably to accept a great number of specific ones, at least initially. (Matsumoto [1987] notes that no less than 76 factors were revealed in Brown's studies [Brown, 1983, 1985b] referred to in Matsumoto [1987, p. 24]). However, if the analysis is to be genuinely useful as research, that is, replicable by other researchers, later analysis may need to reduce these to a

smaller number of more general and more generalisable categories or, alternatively, to focus upon only a sub-set of categories of particular interest.

Crucial to any analysis is that all instances ascribed to a particular category do in fact share the same criterial features. Categories must be formally defined otherwise they will lack any explanatory power and real evaluation of the analysis and possible replication of it will be ruled out. But herein lies the rub. Can the researcher be sure that a strategy or attitude named more than once actually refers to the same thing? There will be problems of definition, particularly with general terms such as motivation. And what is one to make of general comments in the data, that such-and-such is *helpful* or *confusing*, when no supporting information is given? (Retrospective data tends to be full of these interesting yet vague statements which incite a barrage of questions.)

One can only try to be as rigorous and explicit in defining categories as possible. One way is to “go to bed” with the data, to become intimate with it before attempting any categorisation of it, in the hope that categories will suggest themselves. Matsumoto (1987) also suggests that several researchers analyse the data independently before thrashing out a consensus, a helpful though time-consuming approach. A further possibility is to take the data back to the diarist for clarification. However, this gathering of secondary data would invite further editing of the primary data (i.e. it represents retrospections upon retrospections). It would need to be done as soon after the event as possible and that would mean during the period in which the diary was being kept. This intervention, as noted above, would probably influence later entries.

The Contribution of DS

DS have revealed nothing that directly contributes to our understanding of SLA processes (claim 3 is untenable), al-

though they have revealed a wealth of factors that have been perceived by different learners to be important in the enhancement of learning (i.e. claim 2). DS have indicated a great number of learner variables. However, the *participant* studies (as the researcher-as-diarist studies are sometimes referred to) such as Bailey (1980) can only serve to inform us what a particular learner in a particular learning context perceived as being important. Lacking a clear framework of categories of analysis they are impossible to compare. Comparison is also precluded by the range of variables they exhibit as regards the learning context. They therefore lack any explanatory power as the data is not generalisable. As evidence they must be viewed as idiosyncratic, anecdotal accounts (though nonetheless interesting).

Non-participant studies (i.e. involving subjects as diarists) are likely to be more revealing, especially when involving multiple subjects rather than case studies of single learners. But the crucial feature must be frameworks of defined categories that will afford comparisons of studies and replication on other groups of learners where certain variables of learning context may be kept constant. Only then will DS begin to have some explanatory power and without this power the basis for claim 2 is weakened.

Claim 1 is the most tenable of the claims. DS data does raise a great many questions and published DS have all pointed to factors that may be important in language learning. Yet this claim is also weakened unless analyses present defined categories on which hypotheses can be based. Unless what is hypothesised can be tested quantitatively or is at least amenable to support or denial through other qualitative means, the claim remains empty and without meaning. As Chaudron (1986) says:

. . . if we argue that qualitative research serves to generate hypotheses, we must be concerned about the replicability and generalizability of the results. (p. 710)

To its discredit there is a dearth of replication studies in SLA research. To quote Chaudron (1986) again:

Despite many years of qualitative observational studies that should have generated hypotheses about effective teaching and learning behaviors, we have today only a small selection of classroom process variables that can be agreed upon as potentially influential for learning. (p. 711)

Diary Studies in the Research Paradigm

Is there therefore a place for the DS in mainstream classroom SLA research, the foremost aim of which is to describe and explain the nature and processes of SLA? The answer can only be a very weak affirmative. The DS is a limited and exploratory tool which *may* serve to generate hypotheses about learning behaviour and yield information on important learner variables if, and only if, its findings are presented in a framework that is generalisable and accessible to further investigation. However, the charge of lack of generalisability, often made against DS, is perhaps more a reflection of the outcomes of the published studies, still quite small in number, than of inherent weaknesses of the research method itself.

Two Uses for DS

Owing to their exploratory nature, DS may be best used if employed during the initial phase of a research project as a means of throwing up variables to be investigated by other means in a second phase. Alternatively, their use in combination with other research methods may also prove to be fruitful. For example, learners and their teacher might be asked to complete diary entries immediately after the event (i.e. the lesson) followed by interviews to clarify the contents of those entries using a video or audio recording of the event to serve as a prompt. Thus the diary becomes one element of *triangulation*, that is the gathering of data from three distinct sources

that focus upon one event. Although there are some very obvious practical problems involved here, the example is given merely to indicate the possible integration of the DS with other methods of research.

Action Research

There is a completely different and very powerful role for diaries, their role in *action research*. The problems for the DS in mainstream SLA research vanish when they are used in the context of action research, which is research intended to solve immediate problems in classrooms rather than to reveal any general truths about learning. In a number of educational contexts learners have been asked to keep diaries as a means by which teachers can keep in touch with the learners' perceptions of their classroom experiences and as a basis for discussion of problems and of remedial action by the teacher (Hopkins, 1985). It is a powerful tool where learners come to trust it as a form of on-going dialogue with their teacher. It is intended to raise levels of consciousness about learning and to lead to teacher-researcher intervention. Of course there need not be any specific problems to be overcome. Diaries may simply be used as a basis for heightening learners' levels of awareness of their learning. This is recognised by Matsumoto (1987):

Finally, the diary study is not only a research tool, but may also be used for other practical purposes such as self-awareness, self-evaluation, self-improvement, and orientation for other learners – it can be of immediate use for diarist-learners as an aid to their second language learning. (p. 26)

This is true, but one must ensure that the delineation between action research and mainstream SLA research is clearly marked as the DS's strengths in the one become its weaknesses in the other.

Note

1. Classroom SLA here denotes learning in a classroom setting and does not, therefore, imply any distinction between the terms acquisition and learning. The use of procedures to statistically quantify data is the basic feature delineating quantitative from qualitative research methods. Quantitative data includes written or verbal reports from subjects as well as unquantified analyses of transcript data. For detailed discussion see Ochsner (1979), Long (1980), Chaudron (1986), and Henning (1986).

References

- Bailey, K. M. (1980). An introspective analysis of an individual's language learning experience. In R. C. Scarcella & S. D. Krashen (Eds.), *Research in second language acquisition: Selected papers of the Los Angeles second language acquisition research forum* (pp. 58-65). Rowley, MA: Newbury House.
- Chaudron, C. (1986). The interaction of quantitative and qualitative approaches to research: A view of the second language classroom. *TESOL Quarterly*, 20, 709-717.
- Cohen, A. D. (1983). *Researching second language learning strategies: How do we get the information?*, unpublished paper presented at the annual TESOL convention, Toronto.
- Cohen, A. D. & Hosenfeld, C. (1981). Some uses of mentalistic data in second language research. *Language Learning*, 31, 285-313.
- Henning, G. (1986). Quantitative methods in language acquisition research. *TESOL Quarterly*, 20, 701-708.
- Hopkins, D. (1985). *A teacher's guide to classroom research*. Open University Press, Milton Keynes.
- Long, M. H. (1980). Inside the "black box": Methodological issues in classroom research on language learning. *Language Learning*, 30, 1-42.
- Mann, S. J. (1982). Verbal reports as data: A focus on retrospection. In S. Dingwall & S. J. Mann (Eds.), *Methods and problems in doing applied linguistics research* (pp. 87-104). University of Lancaster.
- Matsumoto, K. (1987). Diary studies of second language acquisition: A critical overview. *JALT Journal*, 9, 17-34.
- Ochsner, R. (1979). A poetics of second-language acquisition. *Language Learning*, 29, 53-80.
- Schumann, F. M. & Schumann, J. H. (1977). Diary of a language learner: An introspective study of second language learning. In H. D. Brown, C. A. Yorio, & R. H. Crymes (Eds.), *On TESOL '77: Teaching and learning English as a second language: Trends in research and practice* (pp. 241-249).
- Seliger, H. W. (1983). The language learner as linguist: Of metaphors and realities. *Applied Linguistics*, 4, 179-191.

RELC JOURNAL

A JOURNAL OF LANGUAGE TEACHING AND RESEARCH IN SOUTHEAST ASIA

The RELC Journal: A Journal of Language Teaching and Research in Southeast Asia is one of the professional publications of the Southeast Asian Ministers of Education Organization (SEAMEO) Regional Language Centre (RELC) which is located in Singapore. The main purpose of the journal is to disseminate information and ideas on theories, research, methods and materials related to language learning and teaching. Articles covering a wide range of topics in applied linguistics appear in its pages semiannually in June and December. Special features include information on current research projects in Southeast Asia, book reviews and review articles on topics related to language teaching. In its thirteen years of publication, the journal has reached a wide audience in eighty countries and has been rated as one of the leading journals in applied linguistics throughout the world. Scholars of recognized stature have used it as a means of presenting their findings and obtaining feedback on important issues in language and linguistics.

GUIDELINES

A companion publication, **Guidelines**, provides language teachers with practical and innovative techniques to enliven and enrich teaching in the language classroom. Each issue focuses on a particular theme.

The **Journal** and **Guidelines** are published twice a year and a subscription may begin with the June or December issue. The annual subscription is US\$9.00 for each title. Individual copies may be purchased at US\$6.00.

Please send your order to The Publications Officer
SEAMEO Regional Language Centre
RELC Building
30 Orange Grove Road
Singapore 1025
REPUBLIC OF SINGAPORE