

Ellen Bialystok and Merrill Swain

# Methodological Approaches to Research in Second Language Learning

*The learning of second languages, like any educational undertaking, is subject to such a complexity of influences that the strategies of research cannot afford to confine themselves merely to those inquiry procedures that precedent alone has sanctioned. Bialystok and Swain, laying down three criteria to be used in selecting an appropriate methodology for any intended inquiry, review with examples the respective attributes and limitations of case study research, evaluative research, and experimental studies, by way of establishing that each has its necessary and complementary role in arriving at firm conclusions about language-learning.*

Language learning and language use are influenced by a myriad of factors. These relate to social, affective, cognitive, and other variables which may be observable directly or only through inferential procedures. Because of the complexity associated with language behaviour, the study of it cannot be a simplistic or monolithic enterprise; rather, it must accommodate this complexity by pursuing a variety of courses in order to understand the role of relevant variables.

The progress made in a field depends to a large extent on the existing inquiry procedures. Questions themselves are often formulated as a function of available methodologies rather than as a function of theoretical motivations. Thus, issues that are accessible through current research procedures may take precedence over issues that are less easily subject to empirical observation. Early research in first language acquisition, for example, was dominated by the study of morpheme acquisition and grammatical descriptions (for example, Braine, 1963; Fraser, Bellugi, & Brown, 1963; Brown & Bellugi, 1964), partly for this reason. Similarly, second language research began by pursuing the same course. Techniques from other disciplines which will permit a different perspective on analyses of language learning and acquisition need to be adopted as well. In order to diversify the range of questions, it is necessary to diversify the number and type of experimental procedures.

A proliferation of methodologies, however, increases the danger that inappropriate techniques may be selected for the examination of certain

questions. Each research method is differentially specialized for the investigation of certain problems and certain purposes, and each contributes uniquely to our understanding of language learning in general. An error in selection could produce results which are invalid or non-generalizable.

These considerations also apply within a methodology to the choice of analytical techniques. Demonstration of this point has been recently provided by Rosansky (1976). She assessed the relative merit of using cross-sectional and longitudinal analytical procedures for the study of morpheme acquisition, and found that the longitudinal order obtained for a single subject across a ten-month period did not correlate with the cross-sectional order observed for the same subject at a single point in time. While it is a moot point whether her analysis provided a fair representation of cross-sectional research in that it was based on only one subject, the exercise nevertheless demonstrates the need for caution in selecting the appropriate methodologies and techniques and in stating conclusions about the generalizability of the results.

A similar problem exists with the procedures used within any methodological approach. Larsen-Freeman (1976), for example, compared the order of acquisition of morphemes obtained through different tasks or different elicitation methods. Of the five tasks used — reading, writing, imitating, listening, speaking (Bilingual Syntax Measure) — only speaking and imitating were significantly correlated. For all other measures, different “developmental” orders were observed.

Given, then, the range of alternatives provided by different methodologies, each of which allows a number of different procedures and analytical techniques, on what basis does one choose an appropriate design and consequently avoid the possible ramifications attributable to an error in selection? The decision is based, we suggest, on three criteria.

The first criterion is the statistical assumptions underlying each methodology. Research methods differ largely in the degree of control each imposes on the data. Generally, one aims to find a balance which minimizes the variability in the data and maximizes the generalizability of the findings. These two considerations tend to pull in opposite directions, and each of the methods discussed in this paper resolves the conflict between variability and generalizability differently.

The second consideration is the scope of the study. Language may be examined to any depth or breadth, and the domain considered relevant for a particular study directs to some extent the selection of methodology. So, for example, the study of syntactic development, or negative constructions, or negative morphemes may each be best accomplished through different means.

Finally, the purpose of the study is decisive as well in methodological considerations. Two primary research purposes are the generation of hypotheses and the testing of hypotheses. These are mutually exclusive

to the extent that a single study can adequately accommodate only one of these at a time. The general study of language learning certainly requires both.

In this paper we will outline some of the major methodological approaches used in second language research, indicate some uses for which each is most suitable, and discuss some examples of research in each. The three traditions that will be considered are case studies, evaluation studies, and experimental studies.

### **Case study research**

The case study approach is particularly suited to exploratory inquiry. The data may be examined to any level of detail, and modifications in procedures and analyses may be implemented throughout. This flexibility is a critical feature; it permits the subsequent formulation of specific hypotheses, the generation of which is one of the essential and unique functions of case study research. Flexibility is further evident in the amount of structure imposed on the study; free observation of spontaneous speech in a variety of settings or structured interviews may be used.

The assumption in research of this type is that any individual learner is a microcosm of all learners; all relevant variables may be observed by examining a single instance. Variability between individuals is rarely assessed for its influence on the phenomenon under investigation.

The issues most appropriate for case study are those which examine the development of some aspect of language behaviour in all its contexts. For these studies, the speech of the language learner is documented and descriptive analyses are used to reveal existing trends (e.g. Leopold, 1939). Empirical hypotheses do not generally direct the search through the data, and large-scale parametric statistical techniques are not possible. This is the exploratory use of case study and is most valuable for the initial stages of inquiry.

Analyses such as these have led to the development of descriptive grammars for the linguistic systems of either children acquiring their first or second language (for example, Bloom, 1970; Hakuta, 1976) or adults learning a subsequent one (for example, Rosansky et al., 1974). The danger, however, in establishing descriptive grammar is in the possibility that they become "prescriptive"; when these hypothetical descriptions become the criteria for acceptability, deviations from them may be discounted and potentially important linguistic variability may go unnoticed.

Adaptations of case study paradigms can be used to test specific hypotheses as well. Krashen (1977), for example, used a single case, "P", to test the hypothesis that monitoring occurs under specific condi-

tions. To the extent that the case selected is indeed representative of language learners in general, the procedure is appropriate.

Case study examination is usually longitudinal, but where the target phenomenon is well defined, shorter periods of observation may be sufficient. Peck (1977), for example, observed children in a free situation to accumulate evidence of the existence of language play. The hypothesis was that children engage in an activity in which language is exercised for its own sake rather than for the conveyance of meaning. Observation of several children at play confirmed this speculation. Because the object of inquiry was so specialized, evidence of its existence needed only to be demonstrated for specific situations, that is, play.

Where the data do need to span a longer time period but direct observation is not possible, minimal intervention methods, such as questionnaires or structured interviews, are necessary. In a study by Fröhlich (1976), language learners were questioned intensively about their language learning experiences. Studies such as this place great reliance on subjects' ability to introspect and reflect on their experiences, and are probably susceptible to personal biases and memory inadequacies. Consequently a greater number of subjects than is used for most case study research is required to assure validity.

The primary advantage of case study research is that it allows for the generation of hypotheses. Further, the phenomenon is typically examined in its natural context rather than in some artificial circumstance often created by laboratory research. The problem, however, is to assess the reliability of the observed effect for all language learners. There is no variance in the data attributable to individual differences; the generality of the findings is the question of greater concern. Caution is required in interpreting results of case study research.

## **Evaluative research**

It is largely a concern with language pedagogy that underlies much evaluation research. The necessity for assessing various teaching materials, programs, and methods has resulted in a tradition of evaluation. The general format of the question examined is: What are the effects on achievement of some factor? The factor may relate to aspects of the language program, such as method of teaching, or to characteristics of the learner, such as language learning aptitude.

Extensive testing is usually involved, both to determine the state of the factor under study as well as any achievement that may be directly or indirectly affected by that factor. It is important to notice, however, that the factors and achievement measures collected are carefully tested but not, for the most part, manipulated. Thus, to determine the relationship between aptitude and achievement, measures of both are obtained and compared through procedures such as factor analyses, correlational

analyses, or regression analyses. There is not usually an attempt to modify the variables examined in evaluational studies.

An early evaluative study was conducted by Scherer and Wertheimer (1964) to assess the relative merits of the audio-lingual and grammar translation methods of teaching a second language. Large batteries of tests were administered to groups of students receiving instruction in each method, and comparisons made on various dimensions. Conclusions from studies of this type are concerned primarily with differences attributable to major teaching approaches and are global in scope.

More specific hypotheses may also be tested. An important on-going investigation by Gardner and his associates (for example, Gardner and Lambert, 1972) has been concerned with the relationship between attitude and achievement in a second language. The paradigm again has been to accumulate as much information as possible related to the independent variable, attitude, and observe its effects on achievement across a range of abilities and situations.

Whereas Gardner's research primarily concentrates on the relationship between variables at a single point in time, longitudinal study is also amenable to evaluative research. Primary among these studies is the examination of French immersion programs conducted by Swain and others (for example, Barik and Swain, 1976a). A group of students is followed throughout the program, which is assessed for its effect on French language achievement as well as for its concomitant effects on other cognitive and affective indices.

Evaluative studies typically observe large groups of subjects on a number of measures. These precautions increase the reliability of the findings and make the results generalizable across the population and to other populations. Variance in the data is overcome by the number of subjects involved; deviations from the norm attributable to individual or small group differences are obliterated by the majority response which tends to the norm. The approach, therefore, is particularly suited to studies in which an average estimate of the effect of a factor for a large population is required. Detail about individuals or specific features, such as that possible through case study inquiry, is not typically examined. The possibility exists, however, for the in-depth examination of particular questions (e.g. Barik and Swain, 1976b).

The data analyses used for evaluative research primarily yield information about the relationship between the independent factors examined and achievement. Such results, however, do not resolve the question of causality: the direction of the effect is rarely specified. Arguments could be cited, for example, to suggest that the relationship between attitude and achievement, a reliable finding in the literature, should be interpreted to mean that success in language learning improves one's attitude toward the language and culture, rather than the traditional converse. Clearly the effect is to some extent bidirectional, and information to

evaluate the relative contribution of each process is only tenuously available in the data. Certain independent variables considered in evaluative research are, nevertheless, less open to ambiguous interpretation; a relationship between instructional program and achievement is probably attributable to the program.

Both case studies and evaluation studies are specialized for examining language learning where specific hypotheses may not have been articulated. Case study involves intensive observation of a small number of subjects to determine patterns or features, often without prior evidence of their existence. An examination of the order of morpheme acquisition, for example, need make no *a priori* assumptions about what that order may be; the data, as it were, speak for themselves. Similarly, evaluative research seeks relationships between factors and achievement to determine the variables relevant to language learning. Again, particular outcomes are not necessarily predetermined. Where a specific question is motivating the research, however, an experimental design, in which factors are controlled through intervention of the experimenter, is generally required.

### **Experimental studies**

To test specific hypotheses about language learning, more rigorous experimental designs may be used. Thus, phenomena observed informally in other studies, or hypotheses derived from theoretical considerations, may be investigated to determine specific effects and contingencies.

Experimental paradigms isolate a particular factor by artificially controlling extraneous variables. The experimental tests are designed to reveal patterns which may have been obscured by more global analyses of language learning.

An important concern in experimental studies is the selection of an appropriate sample. The elimination of attention to some of the natural, interacting variables observed in case study research requires that the sample to be tested control for these variables by reflecting a range of individual differences. The assumption is that average scores obtained from a randomly selected group will indicate the norm for the population.

The statistical procedures used with experimental research provide precise information about relationships and causality. In addition to the detail derived from the particular experimental manipulation, more global descriptive effects are also available. For example, an analysis of variance not only indicates the specific behaviour of the dependent variable, but also identifies through the interactions differences between groups attributable to, for example, age, sex, and program. Further descriptive detail may also be abstracted from the data.

Within the constraints of an experimental situation, the manipulation of different aspects of the design allows for diversity in the questions that

may be examined. For example, since experiments are usually conducted at a single point in time, developmental issues are accommodated cross-sectionally. Scovel (1977) measured the development of the ability to recognize foreign accents by testing subjects of different ages. This age difference between subjects constituted the experimental manipulation; all other factors were controlled.

Another application of experimental research is to conduct restricted and specific investigation of questions which, in a broader sense, may be considered evaluative. An assessment of the effectiveness of various treatments, such as teaching materials, may be evaluative if the effects are observed in a large number of naturally occurring settings and subsequently evaluated; the same question for experimental research requires the artificial introduction of the materials in a controlled setting.

A recent study by Bialystok and Fröhlich (1977) demonstrates this point. Students were given French passages to read in one of four cue conditions to assess the efficacy of various cues on reading comprehension. The results showed that comprehension improved when the passage was accompanied by a picture or an English summary statement. The control over other variables in the experiment was such that a single administration of the test produced reliable findings. Evaluative research which is not contrived in this way generally requires repeated examination to assure replicability.

Hypotheses concerning constraints on language learning may be examined experimentally as well. A recent issue in second language learning research is the relative degree to which language learners use explicitly known information as opposed to intuitive information about the target language to formulate utterances in that language. Two hypotheses are that explicit information will be used, first, when there is time to do so (cf. Krashen, 1976), and second, when only a moderate amount of detail is required by the response (Bialystok and Fröhlich, in press). By administering a specifically developed test to second language learners under various conditions of time and detail these hypotheses are currently being examined (Bialystok, in preparation).

The advantage of experimental research is that a great deal of variance is minimized by the exclusion of extraneous variables. Further, specific aspects of language may be examined without the confounding effects of other factors. At the same time, however, this constitutes its greatest danger. The language is examined in an artificial situation and the integration of the findings into a discussion of natural language is sometimes precarious. Generalizability in this case depends largely on the representativeness of the situation from which the data were derived.

The three methodologies, namely case study, evaluation, and experimental, can be mutually facilitative in the examination of specific issues; case study can produce the hypotheses subsequently examined through experimental procedures. Conclusions about language learning cannot be

restricted to information obtained through any one source, as all provide an essential part of the answer to the puzzle.

## REFERENCES

- Barik, H. C. & Swain, M. "Update on French Immersion: The Toronto Study through Grade 3." *Canadian Journal of Education* 1 (1976a): 33-42.
- Barik, H. C. & Swain, M. "A Longitudinal Study of Bilingual and Cognitive Development." *International Journal of Psychology* 11 (1976b): 251-263.
- Bialystok, E. "The Constraints on Explicit Knowledge in Language Comprehension." In preparation.
- Bialystok, E. & Fröhlich, M. "Aspects of Second Language Learning in Classroom Settings." *Working Papers on Bilingualism* 13 (1977): 1-26.
- Bialystok, E. & Fröhlich, M. "The Aural Grammar Test: Description and Implications." *Working Papers on Bilingualism*. In press.
- Bloom, L. *Language Development: Form and Function in Emerging Grammars*. Cambridge, Mass.: The M.I.T. Press, 1970.
- Braine, M. D. S. "On Learning the Grammatical Order of Words." *Psychological Review* 70 (1963): 323-348.
- Brown, R. & Bellugi, U. "Three Processes in the Child's Acquisition of Syntax." *Harvard Educational Review* 34 (1964): 143-151.
- Fraser, C., Bellugi, U. & Brown, R. "Control of Grammar in Imitation, Comprehension, and Production." *Journal of Verbal Learning and Verbal Behaviour* 2 (1963): 121-135.
- Fröhlich, M. "Case Studies of Second Language Learning." Unpublished master's thesis, University of Toronto, 1976.
- Gardner, R. C. & Lambert, W. E. *Attitudes and Motivation in Second Language Learning*. Rowley, Mass.: Newbury House, 1972.
- Hakuta, K. "A Case Study of a Japanese Child Learning English as a Second Language." *Language Learning* 26 (1976): 321-351.
- Krashen, S. D. "Formal and Informal Linguistic Environments in Language Acquisition and Language Learning." *TESOL Quarterly* 10 (1976): 157-168.
- Krashen, S. D. "The Monitor Model for Adult Second Language Performance." In M. Burt, H. Dulay & M. Finocchiaro (eds.), *Viewpoints on English as a Second Language*. New York: Regents, 1977.
- Larsen-Freeman, D. E. "An Explanation for the Morpheme Acquisition Order of Second Language Learners." *Language Learning* 26 (1976): 125-134.
- Leopold, W. F. *Speech Development of a Bilingual Child: A Linguist's Record*. Vol. 1: *Vocabulary Growth in the First Two Years*. Vol. 2: *Sound Learning in the First Two Years*. Vol. 3: *Grammar and General Problems in the First Two Years*. Vol. 4: *Diary from Age Two*. Evanston, Ill.: Northwestern University Press, 1939, 1947, 1949a, 1949b.
- Peck, S. "Language Play in Child Second Language Acquisition." In C. A. Henning (ed.), *Proceedings of the Los Angeles Second Language Research Forum*, June, 1977.
- Rosansky, E. J. "Methods and Morphemes in Second Language Acquisition Research." *Language Learning* 26 (1976): 409-425.
- Rosansky, E. J., Schumann, J. H., & Cancino, H. "Second Language Acquisition: The Negative." Paper presented at the LSA Summer Meeting, Amherst, Mass., 1974.
- Scherer, G. A. C. & Wertheimer, M. A. *A Psycholinguistic Experiment in Foreign Language Teaching*. New York: McGraw-Hill, 1964.
- Scovel, T. "The Ontogeny of the Ability to Recognize Foreign Accents." In C. A. Henning (ed.), *Proceedings of the Los Angeles Second Language Research Forum*, June, 1977.