The Exploring Nature of Methodology in the Current Studies of Language Learning Strategies (LLSs) with Focus on Strategy Inventory for Language Learning (SILL) of Rebecca L. Oxford

Seyed Hossein Fazeli
Department of English Language Teaching, Abadan Branch, Islamic Azad University, Abadan, Iran
Email: fazeli78@yahoo.com

Abstract—There is not a method abandoned as what it might be called a baby-and-bathwater type reaction. Still all the methods have failed to deliver quit the hoped-for miracles. In this way, methodology alone can never be a solution to find what is searching. Rather it is an aid and suggestion. In addition, in the case of methodology, it is true that we are moving in risky area. Research on the related literature of Language Learning Strategies (LLSs) shows that LLSs has a history of only thirty years that is much sporadic (Chamot, 2005a). Recently such strategies have been the focus of specific research (Oxford, 1990), and much of the research was descriptive. The Strategy Inventory for Language Learning(SILL) of Rebecca L. Oxford which is a kind of self-report questionnaire, as an important instrument to measure LLSs, has been used extensively by researchers in many countries, its reliability has been checked in multiple ways, and has been reported as high validity, reliability, and utility(Oxford, 1996a). In the current study, the investigator aims to explore nature of methodology and the use of SILL in the studies of LLSs.

Index Terms—language learning strategies, self-report questionnaires, strategy inventory for language learning, methodology in the studies of language learning strategies

I. INTRODUCTION

One of the human nature characteristics is interesting in research on the related human characteristics. Such research includes two types of approaches that are qualitative and quantitative. The quantitative studies include experimental and non-experimental designs. Each of these approaches and designs has some base of theoretical framework, particular applications, and special methodology. In quantitative research, the goal is to explore the relationship between some variables in terms of research questions or hypotheses.

A literature of research was conducted to identity published and unpublished related sources. The related literature of Language Learning Strategies (LLSs) shows that there are several methods of gathering data. These methods include interview (O’Malley, Chamot, Stewnner-Manzanares, Kupper & Russo, 1985;Wenden1986), think-aloud (Anderson & Vandergrift, 1996; Cohen & Hosenfeld,1981), questionnaire (Bialystok, 1981; Ehrman & Oxford,1990; Oxford & Nyikos, 1989;Ramirez,1986), Dairies (Oxford, Lavine, Felkins, Hollaway & Saleh,1996; Rubin,1975; Tyacke & Mendelsohn1986), observation (O’Malley et al.,1985; Stern,1975 ), list (Stern, 1975), and so on.

II. SELF-REPORT QUESTIONNAIRES IN THE STUDIES OF LLSs

In the most studies of LLSs, a self-report questionnaire is chosen as a basic instrument, because it is possible to use this kind of questionnaire to survey a large number of participants in manner that would be practicing almost impossible using any other method. Questionnaire is used as one of the most widely data-elicitation tool (Gou, 2004). In addition, it is used by many researchers (Bialystok, 1981; Ehrman & Oxford, 1990; Oxford & Nyikos, 1989; Ramirez, 1986). High reliable questionnaires are preferred because we can be sure that, we will get the same result each time we measure the same thing. However if questionnaire is not reliable, we do not know if the changes in the scores are due to changes in the person we are measuring or to some type of error in testing process; in such way, the use of questionnaire is not without criticism.

There are three ways of using self-administered questionnaire instruments. Firstly, using the available questionnaire without any change; secondly, using adapted questionnaire; thirdly, developing new questionnaire.

Based on the availability and characteristics, one researcher will choose one of these questionnaire types. If the same questionnaire has tested for its psychological parameters, it is possible to use it as what it is, and without any change. Nevertheless, since in cross-cultural research, some changes are needed in order to avoid some problems. However, there is lack of cultural sensitivity (Kember, Wong & Leung, 1999). Sometimes, because of lack of particular questionnaire, and dominance of particular situation, one researcher develops new questionnaire.
III. STRATEGY INVENTORY FOR LANGUAGE LEARNING (SILL)

Initially SILL used in North America with foreign language learners by different researchers such as Ehrman and Oxford(1989) and Oxford and Nyikos(1989), and it increasingly used in Asia-Pacific region by researchers such as Yang(1999), Hiso and Oxford(2002), and Griffiths(2003).

Two versions of SILL are available. The first one is used with individuals who their native language is English, and they are learners of other languages than English. It is consists of 80 items (Version 5.1). The second one is used with learners of English as a second or foreign language. It is consists of 50 items (Version 7.0).

In the current study, the later one is discussed. SILL includes Memory Strategies(9 items), Cognitive Strategies(14 items), Compensation Strategies(6 items), Metacognitive Strategies(9 items), Affective Strategies(6 items), and Social Strategies(6 items).

The SILL is a structured survey (Oxford, 1990), which according to Oxford and Nyikos (1989) the strategies which are included in SILL are gathered from extensive literature review. In addition, Oxford (1996a) claim in general, SILL reliabilities have been high, and the reliability remains “very acceptable” (Oxford & Bury-Stock, 1995, p.6). Moreover, Green and Oxford (1995) remark that reliability using cronbach alpha ranging from .93 to .95 depending whether the survey is taken in learner’s own language or in target language. Regarding validity, Oxford and Bury-Stock (1995) point out that the all types of validity are very high. Moreover, factor analysis of SILL is confirmed by many studies (Hsiao & Oxford, 2002; Oxford, 1996a; Oxford & Bury-Stock, 1995).In this way, as Ellis (1994) believes Oxford’s taxonomy is possibly the most comprehensive currently available. Several empirical studies have been found moderate intercorrelation between the items of six categories in SILL (Oxford & Ehrman, 1995). Moreover Oxford (1996b) presents “each sub-scale would have an adequate number of items to facilitate more in-depth understanding of the learning strategies” (p.3).

The SILL has been used in various studies to show how much strategy use correlate with various variables such as gender, learning style, proficiency level, task, and culture (Bedell & Oxford,1996;Bruen,2001; Green & Oxford,1995; Oxford, Cho ,Leung & Kim,2004; Nyikos & Oxford,1993; Oxford & Bury-Stock,1995; Wharton,2000). In addition, Chamot (2005b) claims that SILL is a standard measure, and point that is most important, the most descriptive studies are based on the Oxford’s SILL. In this way, SILL allows easy comparison with other studies and helps the researchers to conclude their studies’ results in comparison with the results which were found by the other related studies in the literature. Such results can support each other in the way that it is useful and helpful for suggestion of pedagogical implications.

The SILL has used with learners of English who are native speakers of different languages such as Chinese, French, German, Italian, Japanese, Korean, Spanish, Thai, and Turkish (Oxford, 1990). Moreover, its reliability reported in many studies as high reliability in different languages (Abu Shamis, 2004; Grainger, 1997; Griffiths, 2002; Oxford & Nyikos, 1989; Park, 1997; Sharp, 2008; Szu-Hsin, Ting-Hui & Tzu-Ying, 2006; Yang, 2007).

The SILL is self-scoring, and paper and pencil survey. It is 5-point scale range from “never or almost never” to “always or almost always”. Currently Oxford and her colleagues are developing a task-based questionnaire to complement SILL (Oxford, Cho, Leung & Kim, 2004).

Initially, Oxford adapted a version of Rubin’s direct/indirect distinction. Oxford (1990) argues, “Direct and indirect strategies are equally important and serve to support each other in many ways (p.12).

As Oxford (1990) points out there is thirty minutes of time in order to administrate SILL. The evaluation for the answer is varied between 1 to 5. For the missing answers, if it was possible, the researcher may ask the respondents to answer them. The respondents are asked to respond the items honestly, answer all the items, and mark their response in correct spaces. In addition, there is no need to have formal training in applied linguistics, psychology, or related fields in order to administrate and score this instrument. However, in keeping with the standards for Educational and Psychological Testing, interpretation of SILL requires professional training in psychological testing and measurement such as reliability.

The respondents are asked to choice the statement which how true of them it is. They are told that answer must be in terms of how well the statement describes them. In addition, they are told that there is no right or wrong answer to these statements.

The focus of instructions is on this state that respondents understand what they are supposed to do. The following instructions are given to respondents:

1. This questionnaire should take about some particular duration of time to complete SILL; 2. There is possible to ask any question about the words and statements that they do not understand; 3. This questionnaire is not a test of respondents’ English language proficiency.

IV. ENGLISH LANGUAGE PROFICIENCY TEST IN THE STUDIES OF LLLS

Since generally there is significant impact of English language proficiency on overall academic achievement of students (Fakeye & Ogunsiji,2009), and specifically there is relationship between strategy use and language proficiency (Abu Shmais,2004), therefore because of nature the use of English LLSs, there is need for a general English proficiency
test for determining the proficiency level of participants in English in order to minimize the effect of English language proficiency on the results which are caused from such English language proficiency.

Determining proficiency in language learning for speakers of other languages is not easy endeavor, and has been discussed by experts (Bachman, 1990; Farhady, 1982). In various studies, language performance levels have been based on scores on norm-referenced or criterion-referenced tests (Bremner, 1999; Dereyer & Oxford, 1995; Green & Oxford, 1995), scores decided by teachers (Chamot & El-Dinary, 1990; Chamot & Kupper, 1989) or self-ratings by learners (Wharton, 2000). Even the extensively used Test of English as a Foreign Language (TOEFL), the International English Language Testing System (IELTS), and the others are not universally accepted as a reliable or valid measure of proficiency.

Usually there are some problems in administration of language proficiency test such as time limitation for administration of full parts of language proficiency test, difficulty of permission to administrate full parts of language proficiency test.

It must bear in mind that the classification of levels in proficiency language of participants in one study is as Jafarpour (2001) defines “the percent classification of subjects by the experimental test that corresponds to those by the criterion” (pp. 32-33) (as cited in Golkar & Yamini, 2007).

V. BACKGROUND QUESTIONNAIRE IN THE STUDIES OF LLSS

Based on some indicators such as the parents’ socio-educational background and occupation, sex, age, mother tongue, the general average of results of previous semester; participants are matched as closely as possible for socio-economic background to minimize the effect of social class.

VI. PROCEDURE OF ADAPTATION OF INSTRUMENTS IN THE STUDIES OF LLSS

In adaptation of instruments from one language to another in research works, some problems occur, such as the problem of translation one questionnaire to another language (Perera & Eysenck, 1984).

It must bear in mind that translation/back translation is one of more effective ways to solve the equivalent concepts of translated and original version of one questionnaire (Behring & Law, 2000), and one researcher can be ideal translator if she or he is fluent in target language.

There is need for some linguists, psychologist, or other professional people who are professional in the fields of related to the questionnaires which are supposed to be translated and adapted. Such people must fully proficient in both languages that the translation is supposed to be done between them. In this way, they are asked to check translated version of questionnaires in order to check the consistency with original version of them. Secondly, since such people are professional in related study of questionnaires, they are asked to check the psychometrics of questionnaires. Whenever full agreement among such people is achieved, and pilot study confirms the items of questionnaires, such questionnaires will be administrated in main study.

VII. PILOT STUDY

Before administration of main study, pilot study is needed to be administrated. The pilot study as “A small-scale replica and a rehearsal of the main study” (Riazi, 1999, p.198) is carried out to achieve the following objectives:

a) To check whether the translated items of questionnaires are understood correctly as the original versions; b) To analyze acceptance degree of psychometrics of adapted version in order to use them as valid and reliable instruments in order to measure what one researcher aims to investigate; c) To test suitability of research methods which are used in order to use research instruments; d) Although all the instruments have provided with needed requirements by the researchers who develop them, but there is need to estimate needed time, cost, and the other requirements to administrate main study; e) To get feedback in order to revise questionnaires to balance between spoken and written forms of the translated language. In such way, the participants can feel the meaning of the items so much better.

The sample for pilot study was selected so as it represents entire sample for participants whom asked to participate in main study. Moreover, sample size in pilot study ranges from 20 to bigger of 65 (Hinkin, 1998).

It is better, in pilot study, at each stage, after participants completed the stage, they are engaged in an informal conversation in their mother tongue. The participants are asked to discuss regarding difficulties at that stage. Based on this type of discussion, the researcher will be aware of the comments in details of the participants in the pilot study.

All the suggested comments are written as a particular list of comments. Such list is checked by the researcher to categorize its items in different categories. Moreover such list is discussed with the those people professional in related fields in order to apply them and affect the instruments of the study. All the suggestions of these people were applied up to everybody agrees upon. The revised instruments are shown to some of participants who participated in the pilot study in order to finalize them.

VIII. DATA COLLECTION PROCEDURES

During data collection, it is better that
a) The instruments are administrated during class time and based on the availability of the participants; b) The researcher, herself/himself, administrate the instruments; c) All the participants participate in the study are explained the goals of study by the researcher; d) All the explanation of the materials is performed through mother tongue language of the participants; e) To increase the credibility of the response, the participants are asked that they should be honestly in their answers, and they should not spend too much time on any of the items; f) The respondents are asked to ask any question or doubt it they have; g) After completion of answering the questionnaires, the respondents are asked whether they answer all the items or not; and whether they mark their response in correct spaces or not. In each of both states, if the respondents did not answer all the items or not mark their responses in correct spaces, they are asked to revise those items, and answer those items or mark their responses in correct spaces; h) It must bear in mind that all the participants must be volunteer to participate in all stages of the study.

During data collection, the sample drawn from the population must be representative so as to allow the researchers to make inferences or generalization from sample statistics to population (Maleske, 1995). As Riazi (1999) presents “A question that often plagues the novice the researcher is just how large his sample should been order to conduct an adequate survey or study. There is, of course, no clear-cut answer” (pp.242-243). If sample size is too small, it is difficult to have reliable answer to the research questions. If sample is too large, it is difficulty of doing research. Moreover, number of participants may affect the appropriateness of particular tool (Cohen & Scott, 1996).

IX. RATIONALE FOR THE METHODS OF DATA COLLECTION

There are various methods to collect data, such as survey tools and written questionnaire (Gu & Johnson, 1996; Fan, 2003), interview (Gu 2003b; Parks & Raymond, 2004), think-aloud or verbal reports (Anderson & Vandergrift, 1996; Goh, 1998; Nassaji, 2003), diaries or dialogue journal (Carson & Longhini, 2002; Oxford, Lavine, Felkins, Hollaway & Saleh, 1996), recollective narratives (Oxford, Lavine, Felkins, Hollaway & Saleh, 1996). Such measurements are used in the single form of method (separately) or as component methods (single set of methods) based on nature and goals of research works. For example, Griffiths (2004) used self-report (SILL) and interview in order to find the relationship between LLSs and proficiency in her research work. Rubin (1975) did by means of observing students in classrooms, talking to good language learners, and eliciting observation from teachers. However, one of the main difficulties in the study of LLSs is direct observation of LLSs (Griffiths, 2004).

Every one of these methods has specific structure and characteristics. For instance, according to Nunan (1992), interview vary from structured, through semi-structured to un-structured, and as Naiman, Frohlich, Stern and Todesco (1978) claim “the interview proved to be a useful research technique” (p.35). Moreover as O’Malley et al. (1985) state that “generally we had considerable success in identify learning strategies through interview” (p.35). Alternatively, the multiple-methods procedure has specific characteristics. For instance, it is used whenever the time is available. When the time is restricted, researcher should use the most reliable and valid strategy assessment measure (Oxford, 1996a). In this way, Cohen (1998), McDonough (1995), O’Malley and Chamot (1990) claim that each kind of data collection has its own limitation.

One of important methods is self-report method. Self-report is always subject to errors. For example since usually some of the strategy items are vague, and therefore open to differing interpretation, thereby possibly affecting reliability, it causes to the results, as Gu, Wen and Wu (1995) state that such results can be “dangerously inadequate and unreliable” (p.7). Nevertheless, there is not better method has been suggested to identify learner’s mental process. In addition, as Dornyei (2003) points out advantages for self-report, as versatility, cost effectiveness, and efficiency in terms of staff and students time and effort. Moreover, most language research has favored the use of questionnaire, inventories, and survey because data can easily be collected by a large number of participants and analysis is uncomplicated (Cohen, 1998; Oxford, 1990, 1996b). Self-administered questionnaire may have much less interview bias (Gorard, 2001). In addition, as Cohen (1987) claims that LLSs are internal mental processes and not directly observable behavior. Their identification and description have relied greatly on learners’ reports (as cited in Chamot & El-Dinary, 1999).

Although most research has employed multi-data collection for gathering and validating learning strategies (Ellis, 1994), and in order to amplify the quantitative questionnaire data, it is need to employ complementary qualitative data collection, as recommended by Chaudron (1986), but as O’Malley and Chamot (1990) point out the results of data collection procedure varied considerably when there are multi-data collection procedures.

At finally yet importantly, it must bear in mind that it is the best when one method is dictated by nature of the problem rather than academic fashion and prestige.

X. DATA ANALYSIS PROCEDURES

In the most studies, the data obtained through the instruments is entered into database (Statistical Package for Social Sciences (SPSS) software) to enable data analysis to be carried out.

In the analysis of data, it must description of the entire sample of participants, finding reliability and validity of instruments, correlation analysis between the independent variables and the dependent variables, checking of prediction of some variables through the other variables, and so on.
Data analysis includes different data analysis procedures. One of these procedures includes calculating descriptive statistics. Such descriptive statistics including range and mean of age of participants, reported frequency of strategy use, reliability of instruments, and so on. One of important issues in descriptive statistics is issue of reliability of the instruments for entire group. Internal consistency and test-retest reliability are the most commonly used methods to check the reliability of tests and measures (Costa & McCare, 1992). Moreover “Cronbach Alpha, a measure of internal consistency, was chosen as the most appropriate reliability index” (Oxford, 1996a, p.31), and it is identified to be an appropriate measure which is applied in order to estimate the reliability of items within one instrument (Pedhazur & Schemelkin, 1991).

The second procedure of data analysis includes Pearson Correlations that used to identify the strength and direction of the relationship between variables. As known to the researchers in the field, correlation does not imply causality, but it does provide a picture of relationships. Moreover, Pearson correlation cannot be used when one of two variables is nominal.

The third procedure of data analysis includes Analysis of Variance (ANOVA) that is an analytic tool. In non-experimental research, ANOVA does not show the same meaning as experimental research. In non-experimental research, ANOVA does not mean causality between the independent variables and the dependent variables when there is significant relationship. In this way, the use of ANOVA in non-experimental research is criticized if the goal is finding casual relationships (Johnson, 2001). Moreover the use of ANOVA in non-experimental is perfectly acceptable when the goal is not causality according to top statisticians (e.g. Johnsen, 2001). In addition, ANOVA has been frequently used for many years in non-experimental research (Johnson, 2001).

ANOVA or F-test is used as a single or one-way, as two-way or as N-way analysis in order to compare more than two variables which are studied (Riazi, 1999). Such comparison includes means (Mousavi, 1999). Depend on the type of research whether experimental or non-experimental research, the goal of ANOVA is varied. In such way, correlation is used to find degree and direction of relationship between variables, and ANOVA test the significance of relationship.

The fourth procedure of data analysis includes multiple regression analysis. As Newton and Rudestan (1999) point out it is used to find relationship between multiple distributed independent variables and a single dependent variable. As Mousavi (1999) remarks regression analysis is “a statistical technique for estimating or predicting a value for a dependent variable from a set of independent variables”(p.320).

The fifth procedure of data analysis includes factor analysis in order to determine whether it is possible to divide the items of the questionnaire into sub-groups. As Riazi (1999) remarks factor analysis is “A statistical procedure for analyzing the intercorrelation among a number of measures that reduces the set to a smaller number of underlying factors” (p.97). In such procedure, it is assumed that there are probably one or more underlying factors (Mousavi, 1999). In addition, some more procedure of data analysis, exploratory factor analysis, is performed. Such factors are used for the adaptation of instrument or the development new instruments.

XI. LIMITATIONS OF METHODOLOGY IN THE STUDIES OF LLSs

All the education quasi-research deal with living human beings are occur out of laboratory conditions have limitations (Gall, Gall & Borg, 2003). All the studies have several of limitations of methodology are ones common in the literature. Firstly, the need for a large scale.

Secondly, it is exclusive reliance on self-report responses to the questionnaires. Since the questionnaire is self-report and single source of information, it is not clear whether participants actively used the strategies, they indicated. Their response may not be just their beliefs and thoughts that they have about their use of strategies. In order to investigate students’ actual use of strategies, the researchers must observe classes, use think-aloud procedure (introspection), interview, and so forth. Moreover, there may also have been some unclear points in questionnaires themselves. In addition, the vagueness of wording has been another persistent problem in using questionnaire (Gu, Wen & Wu, 1995). Another difficulty in cross-language research involves translation. In the case of SILL, SILL does not describe in detail the LLSs a student uses in responses to any specific language task.

The third one, there is an issue in the statistical procedures. The reliability estimates internal consistency may not be appropriate to measure something that could fluctuate in short period; however, it is popular in the studies of LLSs. The test-retest reliability measure is better indicator of reliability in this type of research.

The fourth issue, since measurements which are developed in the western countries may not be so successfully employed in the eastern countries, and many value measurements which are developed in western countries are not success to assess in eastern countries (Matthews, 2000; Schwartz, Malech, Lehmann, Burgess, Harris & Owens, 2001). In the case of used instruments, some limitations may disappear. Such limitations are as characteristics of cross-cultural research and its instruments.

XII. CONCLUSION
Research on the related literature of LLSs, shows that LLSs has a history of only thirty years which is much sporadic (Chamot, 2005a). Recently such strategies have been the focus of specific research (Oxford, 1990), and much of the research was descriptive.

Although each study employs some interesting methodology, but because of uncontrolled extraneous variables, certainly it is not conclusive and comprehensive results in nature. And it is rarely possible to adequately control for all variables in any natural research. In this way it is better that, it should be some research methods to corroboration results of each others.

There is need to have more psychometrics of instruments as possible which are used in study. For example, regarding questionnaires, the reliability must be on the threshold of .70 (de Vaus, 1995), but this is of instruments in any natural research. The value of .70 is high enough to use the results of the research (Chamot, 2005b). In this way, it must so much attention about this matter that the reliability must be more higher as possible and it must tested in reliable way.

REFERENCES


Seyed Hossein Fazeli is pursuing Ph.D. program in linguistics. He has received his M.A. degree in linguistics. His interest includes Language Learning Strategies. Currently he is as a lecturer at Department of English Language Teaching, Abadan Branch, Islamic Azad University, Abadan, Iran.