Adaptation of the Strategy Inventory for Language Learning (SILL) for students aged 12-15 into Greek:
Developing an adaptation protocol

Zoe Gavriilidou & Lydia Mitits

Democritus University of Thrace
zoe@otenet.gr, lydiamitits@gmail.com

Abstract
The purpose of the study was to develop an adaptation protocol for Oxford’s Strategy Inventory for Language learning (SILL) from English into Greek to be administered to monolingual/multilingual students aged 12-15 in secondary schools in Thrace, Greece. This study focuses on following the appropriate adaptation protocol in order to maximize the questionnaire reliability and validity, both when used with the particular learners and when used to compare scores across cultures and languages. The original scale was translated into Greek, back-translated and reviewed. Cross-cultural adaptation included the experts’ revision, followed by the instrument administration to 50 participants. Its internal consistency was .91. Test-retest reliability ranged from fair to good for the total scale and its six-subscles.

Keywords: questionnaire adaptation, SILL, language learning strategies, reliability, validity

1. Introduction
The last decade has seen a growing interest in studying language learning strategies in Greece. The works of Kazamia (2003), Gavriilidou & Papanis (2007; 2010), Gavriilidou & Psaltou-Joycey (2010), Psaltou-Joycey (2010), Vrettou (2011) and others investigate ways of identifying and measuring strategies used when learning a foreign language. That research has identified the problem of not having a valid and reliable instrument for measuring language learning strategy use and has demonstrated the need for a relevant instrument adaptation.

---

1 This study is part of the Thales project MIS 379335. It was held in the frame of the National Strategic Reference Frame (E.Σ.Π.Α) and was co-funded by resources of the European Union (European Social Fund) and national resources.
2. Literature review

2.1 Literature on questionnaire adaptation

There are advantages of adapting an instrument rather than constructing a new one. The process is cheaper and faster since development, validation and norming of a new instrument are both expensive and time-consuming, and the database that is created allows validity studies of the adapted questionnaire as well as cross-linguistic and cross-national comparability (Hambleton & Patsula 1998). However, errors occur during an instrument adaptation in the area of cultural/language differences, technical methods and the way the results are interpreted (ibid.). For instance, there are errors resulting from not establishing construct validity, those related to the inappropriate item formats and translators, and those stemming from wrong interpretation of questionnaire scores. Even though there is often the assumption that the translation of an instrument will retain the psychometric properties such as validity and reliability of a scale, Hambleton & Patsula (1998) point out the difference between a questionnaire translation and adaptation. According to them, translation is just a step in the process of adaptation, which also includes the following: deciding whether an instrument contains construct validity across cultures and is appropriate to use, selecting translators, deciding on the procedure and, finally, adapting the questionnaire and checking its equivalence at the level of content.

Another important issue when adapting questionnaires from different cultural and language backgrounds is the bias that may occur. Bias and its counterpart, equivalence, are two essential concepts in instrument translation and adaptation. There are three types of bias: construct, method and individual items bias. Construct bias occurs when the construct measured is not identical across cultural groups. Method bias refers to incompatibility of samples, e.g. if cultural groups have different educational background, different levels of motivation or interest in the instrument completion as well as ambiguous test instructions and administration problems. Item bias is seen as a distortion of meaning at the item level, when biased items have different meanings in different languages and cultures (van de Vijver & Leung 1997a, 1997b). van de Vijver & Tanzer (1997) propose the following in order to overcome bias: the use of informants with expertise in local culture and language; the use of samples of bilingual subjects; the use of test-retest, training and/or intervention studies; linguistic and conceptual item bias detection and, finally, psychometric methods of item bias detection.
Since there is a growing interest in using questionnaires in a specific linguistic and cultural setting, there is a need to set standards of how those questionnaires should be adapted to allow use in different cultures and languages without compromising the instrument’s reliability and validity. What is important is to “reach equivalence between the original source and target versions of the questionnaire” (Beaton et al. 2000: 3186). According to the International Test Commission Guidelines for Translating and Adapting Tests (2010), there are four issues which have to be considered when a questionnaire is to be adapted: context (specific information on how socio-cultural and ecological contexts might affect scores should be offered), adaptation (the adaptation process should take full account of linguistic and cultural differences of the target population), administration (questionnaire administration instructions should be in the source and target languages to minimize the influence of unwanted variation across populations and score interpretation), and score interpretation (appropriate statistical techniques to establish the equivalence of the different versions of the instrument and identification of problematic components should be applied and documentation of the changes should be provided).

For the purposes of the present study the adaptation process was broken down into three steps: the translation process, the cross-cultural verification and adaptation, and the verification of the psychometric properties of the questionnaire (Rahman et al. 2003). The translation process included: initial translation by two independent translators, synthesis of the translations during which any discrepancies between the two initial translations were resolved and back translation into the original language. The cross-cultural verification involved expert committee review which helped achieve semantic, idiomatic, experiential and conceptual equivalence, while psychometric properties of the instrument were verified by pretesting the final version. Finally, final reports drown for all the stages were submitted to the coordinating committee (Beaton et al. 2000).

The questionnaire reliability (e.g. internal consistency and test-retest), validity evidence, test bias, test administration procedures, and test-takers variables that may influence validity and interpretation of results are considered in the final stage of the adaptation process (Turner et al. 2001). Beaton et al. (2000) maintain that a careful cross-cultural adaptation should ensure content and face validity between the source and target versions of the scale. In other words, if the original scale is reliable and valid so should be the adapted one. As this may not always be the case on account of
subtle cultural differences, psychometric measurements should be employed in order to ensure statistical or psychometric properties of a questionnaire.

Bearing all of the above theoretical considerations in mind, the SILL questionnaire version 7.0 (ESL/EFL) was selected for its content validity and reliability as it has been used in different language and cultural settings. It can be applied to diverse populations after cultural adaptation. Therefore, in order to study the strategic profile of Greek students, it was deemed more feasible to use a tried and tested instrument after appropriate adaptation than to develop a new one.

### 2.2 Literature on the SILL adaptation

The SILL 7.0 (ESL/EFL) for learners of English as a second/foreign language (50 items) was developed by Oxford (1990). The scale consists of 6 subscales: 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) and uses a choice of five Likert-scale responses (1-5) for each strategy described: from “never or almost never true of me” to “always or almost always true of me”. Oxford & Ehrman (1995) point out that although the current ESL/EFL SILL was constructed using six subscales, reliability of the SILL is determined with the whole instrument. This is because the six subscales are strongly correlated with the SILL mean (.66 to .81) and moderately correlated with each other (.35 to .61). In general, the ESL/EFL SILL reliabilities reported in the literature have been high.

The SILL has been translated into at least 17 languages and administered to 10,000 learners approximately (Chamot 2001). The majority of those learners have been native speakers of Spanish, French, Chinese, Japanese, etc. As far as the SILL reliability after linguistic and cultural adaptation is concerned, Oxford (1996) lists a number of research results which prove its high reliability when translated into a native language of the respondents and then administered. In general, the translated versions of the SILL have had high reliability index expressed through Cronbach’s alpha which varied between .91 and .95. Indicatively, we mention the Chinese translation (Yang 1992), Japanese translation (Watanabe 1990), Korean translation (Oh 1992), and Turkish translation (Demirel 2009).

Although the SILL is a standardized measure with versions in many different languages which can be used to gather and analyze information on large number of language learners, it has also received some criticism. LoCastro (1994; 1995) in
Macaro (2006) argues that language learning inventories, such as the SILL, lack validity on account of the fact that they are not transferable across sociocultural domains. It will be argued in this paper that sociocultural bias can be overcome if a detailed adaptation procedure is employed. Oxford (1996) supports that the SILL construct validity is represented in the relationship between the questionnaire and the language performance, meaning that, generally, more advanced learners use more strategies more frequently. Construct validity of the SILL has also been studied in relation to the ESL/EFL setting, learning styles, gender, motivation, etc. and it has been found that there is a strong relationship between the SILL score and the aforementioned independent variables (Oxford, 1996).

When it comes to adapting the SILL into Greek, there have been two relevant studies so far. One focuses on measuring the frequency of language learning strategy use in adult Greek learners of English (Kazamia 2003), while the other records the frequency of use in primary school children who are learning English at school (Vrettou 2011). Both studies use adapted versions of the instrument developed by the researchers themselves and they both contain elements of a thorough adaptation process into Greek. However, those adapted versions have been developed to cater for adult learners and primary school children respectively and not for adolescent learners aged 12 to 15. The present study attempts to fill this gap by developing a carefully organized questionnaire adaptation and particularly of SILL for secondary level adolescents in Greek.

3. Research method

3.1 Purpose of the study
The purpose of the study was to develop the process of adapting the SILL from English into Greek in order to administer it to monolingual and multilingual students in junior high schools in Thrace, with the view to establishing the appropriateness of the adaptation process at the pilot stage. A further adaptation will be carried out as part of the project of profiling the language learning strategy use of students in Greek primary and secondary schools.

3.2 The context and participants
For the purposes of the main research phase, it was deemed necessary to pilot the adaptation of the SILL for the particular learner population. The participants in our
study come from various social and cultural backgrounds, as there are many of them with L1 other than Greek and belong either to the Muslim minority in Thrace, which is either Turkish-speaking or Pomak-speaking and in fewer cases Romani-speaking, or to immigrant families from former eastern bloc countries. The most distinguishing and, at the same time, demanding feature of our target population is its diversity with respect to its linguistic and cultural backgrounds. The majority of the participants belongs to a L1 Greek-speaking homogenous group and is not expected to encounter any particular difficulties during the questionnaire administration. Another large group is L1 Turkish-speaking participants who are characterized by certain idiosyncrasies. While the primary school children, who are taught in Muslim minority schools and find themselves in a homogenous environment, do not object to answering the Turkish version of the questionnaire, the secondary school students studying in mainstream heterogeneous learning environments do not appreciate being segregated by language and insist on answering in Greek. During the entire translation protocol particular attention was paid to providing sufficient comprehension levels for this latter group of respondents who have L1 other than Greek and thus may need some simplification in language but not in concepts.

Moreover, Greek is the only common language for all the participants who have at least reached the intermediate (B2) level of proficiency on the CEFR (2001), meaning that they should not have particular difficulties in responding while their English proficiency levels are mixed so they could not be expected to respond to the SILL in English with a high degree of accuracy. No attempt was made to translate the SILL into the learners’ first languages since it was infeasible to establish if they are literate in their L1 and to what level (except for the Muslim group with L1 Turkish who receive education in two languages, Greek and Turkish).

### 3.3 Adaptation protocol

The process of adaptation was broken down into three stages: (a) the translation process, (b) the cross-cultural verification and adaptation, and (c) the verification of the psychometric properties of the instrument. The translation process consisted of the initial translations, synthesis of the translations and back translation. The second step included the expert committee review in the light of the focus group suggestions and other verification methods. Finally, in the third stage, the questionnaire was administered and its psychometric properties were verified (see figure 1). In the
following section we will describe the steps taken and the changes made in each stage of the adaptation process in detail.

Figure 1. Adaptation protocol

4. Results
4.1 The translation process
The process of translating the SILL from English into Greek took place at three levels: linguistic /semantic, technical and conceptual as proposed by Rahman et al. (2003), and equivalence between the original and translated versions was considered at each level. To these three, the 'comprehension level' was added to ensure that the target population understood the translated material as easily as the source population for whom the original questionnaire was designed.

The initial translation was undertaken by two translators who have an excellent command over technical and colloquial aspects of both English and Greek as well as an in-depth insight of the cultures. The first translator was an ‘informed’ translator, qualified in the area investigated by the questionnaire and with necessary technical and scientific background in order to understand the concepts and constructs used. The second one, the ‘uninformed’ translator, was not informed about the concepts measured and did not have any particular knowledge of the subject matter. Both produced written reports with their comments on the difficulties they experienced. Next, they compared their versions and synthesized a new one while reporting the process of the synthesis. Both agreed that there were no particular linguistic and
Adaptation of the Strategy Inventory for Language Learning (SILL)

semantic issues to be resolved apart from certain items which demanded careful paraphrasing as literal translation would lead either to ambiguity or misunderstanding of the concepts in question. Such items were the following: ‘I physically act out new English words’ (item No7) or ‘I try to find patterns in English’ (item No 20) which were paraphrased in Greek. The translators’ reports made it evident that a number of items could cause cultural bias and it was the responsibility of the panel of experts to remove that bias and review the translation.

An English teacher (a native speaker of English) and another English native speaker then back translated the questionnaire into English. Both produced blind back translations during which the back-translators were not informed about the concepts under investigation. This process enhanced content validity of individual items as it ensured a consistent translation. The back-translators’ written reports revealed that all items contained the same concepts as the original ones and there was no need for revision after the back-translation.

4.2 Cross-cultural verification and adaptation

The items were discussed with 2 Greek language teachers and 2 English language teachers in the junior high schools that the students in question attended, 1 Russian L1 speaking teacher of Greek and 1 Turkish L1 speaking teacher of Turkish, both of whom were Greek university graduates. The particular key informants were selected on the grounds of their profession, knowledge of the languages in question and familiarity with the student population. These key informants were given the translated questionnaire and asked to comment on each item, especially those that had proven problematic in the first step of translation. They agreed that the translation was generally easy to understand and that the students would not have any particular difficulties in comprehending the linguistic and syntactic level. The most objected item was item No 43 ‘I write down my feelings in a language learning diary’ as the concept of keeping a diary in order to record one’s feelings about learning a language is simply not experienced in Greek education. This item was removed. Another comment referred to the technical issue of the questionnaire format, layout and rubrics in order to make it more reader-friendly, less overwhelming and intimidating for teenagers. Thus care was taken in order to reduce method bias, in particular administration bias which is found in the ambiguous instructions for participants and
the guidelines for administrators. In our case it was overcome by the adaptation of the layout and provision of a detailed manual and administration protocol.

The questionnaire was given to a convenience group of 8 students from the study population. The participants were exposed to the questionnaire in pairs and were asked to note down problems of comprehension, language and cultural relevance and were encouraged to give suggestions which led to the second revision of the translation.

To eliminate any comprehension difficulties this second revision was further administered to 30 12-year-old students with L1 Turkish as that particular target group was expected to encounter most problems due to the fact that their L2 Greek proficiency levels are mixed, ranging from low to intermediate. The focus groups' remarks were recorded and transcribed. The written report was submitted to the panel of experts comprising the two researchers and methodology designers, a professor of applied linguistics, and the four translators. Discrepancies were removed, differences were discussed and the seriously disputed items were changed. 31 items were straightforward, and no further major changes were applied to them. The remaining items were modified at later stages of the procedure.

There is a general agreement between the original English version and the Greek translation. Two major alterations were made and the first included memory strategy item No 4: ‘I remember a new English word by making a mental picture of a situation in which the word might be used’ which caused problems on the conceptual and comprehension levels, as all the subjects asked for clarifications and still could not understand the notion, probably because of their age and level of cognition. Since this item checks mental learning processes as well as learning style preferences (visual type learners) as does item No 9: ‘I remember new English words or phrases by remembering their location on the page, on the board, or on a street sign’, it was not deemed essential for the purposes of the questionnaire administration in the present study and thus was removed. The strategy No 31 in the SILL version 5.1: ‘I use reference materials such as glossaries or dictionaries to help me use the new language’ was used instead and added to the cognitive strategy category. There is a general agreement in the literature that this particular strategy is significant in second/foreign language learning and is included in a number of strategy lists. In Greece it was employed in questionnaires adapted to record language learning
strategy use in adults and primary school children respectively. (Kazamia 2003; Vrettou 2011).

Another alteration included the substitution of affective strategy item No 43: ‘I write down my feelings in a language learning diary.’ with another affective strategy item No 67 from Oxford’s SILL version 5.1: ‘I actively encourage myself to take wise risks in language learning, such as guessing meaning or trying to speak, even though I might make some mistakes.’ It was made as a result of the focus group comments and key informants’ suggestions which led us to conclude that the particular item (No 43) is invalid since none of the subject reported using it and they seemed confused by the concept of keeping a language learning diary. On the other hand, strategy item No 67 was considered extremely important by the experts because it formed a crucial part of the research hypothesis on multilingual language learners for whom the SILL is being adapted.

The remaining modifications were slight and for reasons of better comprehension. The rest of the adaptations had to do with retaining the linguistic or semantic equivalence of similar meanings, making sure that the translated meanings remain as close as possible to the original ones while obtaining an identical meaning of concepts which may have different cultural understandings.

4.3 Verification of the psychometric properties of the instrument

The psychometric qualities of the Greek adaptation of the SILL are presented in terms of validity and reliability. According to the written reports submitted by the panel of experts, it can be argued that the Greek version of the questionnaire is as valid as the original one concerning the item-level equivalence since the careful adaptation procedure has ensured semantic, idiomatic, experiential and conceptual equivalence. Its validity is further improved by resolving technical issues of questionnaire translation. The adapted SILL will be further tested for its content validity through confirmatory factor analysis, where a six-factor model based on the six subscales suggested by Oxford is constructed and tested (see Demirel 2009). This statistical analysis method will be employed in the main stage of the research for which the adaptation protocol was developed and where the participants will surpass 2000 students.

To check the SILL’s internal consistency a Cronbach’s Alpha analysis was performed. To check the stability of SILL scores over time, test-retest data were used
and the intra-class correlation coefficient was computed using SPSS version 15. Reliability analysis was performed in October 2011. It involved 25 L1 Turkish and 25 L1 Greek speakers, second year junior high school students. In order to measure test-retest reliability, the scale was re-administered to the same participants after three weeks. All experimental procedures were approved by the Institutional Review Board for investigations involving human participants. Written informed consent was obtained from the legal guardians of the participants before they were allowed to participate in the study.

The internal consistency coefficient for the whole scale was calculated and Cronbach’s alpha was found at .91 suggesting a high degree of internal consistency of the SILL. Cronbach’s alpha for memory strategies was .71, for cognitive strategies was .82, for compensation strategies was .51, for metacognitive strategies was .48, for affective strategies was .78 and for social strategies was .82.

Test-retest reliability for the total scale and the sub-scales ranged from fair to good (Total scale: r= .778, p<001, Memory strategies r= .831, p<001, Cognitive strategies: r= .874, p<001, Compensation strategies: r= .761, p<001 , Metacognitive strategies: r= .696, p<001, Affective strategies: r= .851, p<001, Social strategies r= .861, p<001 ) indicating that at least within the time frame considered here scores of SILL mirror stable individual differences.

5. Discussion

The aim of the study was to develop the process of the SILL adaptation, to pilot it and to establish its validity and reliability in order to follow the same adaptation process in the main stage of the research. Criteria were set and met during the SILL adaptation. Linguistic and cultural differences were taken full account of and appropriate statistical techniques were applied. Reliability of the questionnaire was measured and found to be sufficiently high and total scores on the SILL were reliable over a three week interval. Validity was discussed with respect to its construct across cultures and languages, but also during focus groups and panel of experts meetings and with respect to other relevant studies carried out in the Greek context. For example, a completely invalid item checking the use of a language learning diary (Affective strategy item No 43) was spotted and excluded as it had lowered the validity in two previous studies in Greece (see Kazamia 2003; Vrettou 2011).
Every attempt was made to reduce the bias that occurs during translation. Construct and item bias was noticed and dealt with in order to overcome the problem of measuring different constructs in different cultures or distorting the meaning of individual items. Method bias, in particular administration bias discovered in the ambiguous instructions for test-takers and guidelines for administrators, was overcome by the adaptation of the layout and provision of a detailed manual and administration protocol.

As proposed in the literature, the present study used expert informants concerning Greek, Turkish and Russian language with experience in teaching in the Greek educational context. It also used representative samples of the research population which provided significant feedback on the linguistic, technical and conceptual levels of the adapted instrument.

The researchers argue that, since the translation protocol was carefully carried out, socio-cultural bias should be avoided and the results should be reliable and as Oxford claims: “The SILL can be administered in the respondent’s native language or a foreign or second language with confidence that measurement error is minimal.” (Oxford 1996: 32)

The process, during which items were translated and in some cases replaced in order to make the instrument more reliable and valid in Greek and with a culturally diverse student population, has been described. But it is not only the item equivalence that was adapted. Other methods, such as the use of key informants, panel of experts or audio transcriptions, allowed for the translation and adaptation to provide a valid measure of language learning strategies used by teenage students in Greece since it would be wrong to assume that the same instrument once translated will be equivalent per se in the new linguistic and cultural context.

6. Conclusion and further investigation
It can be concluded that the process of adapting the SILL from English into Greek recorded in this paper, however time consuming and costly, is the most effective way to produce an instrument for measuring the frequency of language learning strategy use of early adolescent language learners of various linguistic and cultural backgrounds who receive formal education in Greek junior high schools. It also allows for comparison of data and findings across nations as it provides the opportunity to examine language learning strategies of those for whom there was
previously no translated version of the SILL. The carefully planned and executed adaptation process ensures high instrument reliability and validity and offers other researchers interested in questionnaire adaptation a procedure that overcomes most of the problems entailed when instruments are used in different languages and cultures. Further investigation will produce a great amount of data to be collected from language learners enabling the researchers to submit the instrument to more psychometric measurements thus adding to reliability and validity of the adapted SILL questionnaire.

References


