Relative Importance of Learning Variables on L2 Performance[*]

Gabriela Olivares-Cuhat (University of Northern Iowa)

Abstract

In this article, the results of a pilot study investigating the relative importance of various learner variables on L2 performance are presented. The study was conducted with the participation of forty students enrolled in a beginning Spanish class at a large midwestern university. The CANAL-S test, FLCAS and SILL surveys were administered to evaluate language learning aptitude, anxiety and learning strategy use. Demographic data and exam grades were also recorded. The data was first analyzed within the framework of a causal model with correlational and multiple regression analyses. A second step involved the evaluation of the relative importance of the learner variables. Third, the validity of the causality postulated between foreign language anxiety and L2 performance was examined. Overall, this paper confirmed the prevailing role played by language learning aptitude and anxiety, and demonstrated the applicability of a multivariate model to analyze the impact of individual differences on L2 performance.

1 Introduction

Since Carroll and Sapon's groundbreaking work on aptitude (1959), many learner variables have been proposed to account for individual differences found in language learning. Today, these various attributes are grouped into general categories, including cognitive (e.g. language aptitude and learning styles), affective (e.g. attitudes, motivation and learning anxiety), metacognitive (e.g. use of language learning strategies) and demographic factors (Horwitz 2008). Now, in spite of the extensive body of research found on these topics, only few studies have attempted to explain how different types of learner variables interact and compete with one another in the L2 classroom. As a result, language teachers who use a learning variable perspective to shape the content of their courses will find it difficult to answer the question of ranking these various factors in terms of their influence on academic performance.

Therefore, the primary goal of this paper is to examine the relationship between some major language learner variables from the point of view of their relative importance on L2 performance. To this end, a pilot study was conducted at a large midwestern university within the framework of a beginning Spanish program.

A critical step of the method presented in this article is to posit a causal link between learner variables and L2 achievement. However, it has been suggested that a spurious relationship with L1 language skills may invalidate this hypothesis in the case of language learning anxiety (Sparks/Ganschow 1995, 2007). Consequently, a second goal of this pilot study is to assess the robustness of the assumption of a cause-and-effect relationship between language learning anxiety and L2 performance.

2   Review of previous research

2.1   Foreign language aptitude

Foreign language aptitude, which may be broadly defined as "language learning ability" (Dörnyei 2005: 32), has been the most studied learner variable and remains an important area of research (Bialystok/Fröhlich 1978; Skehan 1989). The first cognitive test to effectively predict language learning success, the Modern Language Aptitude Test (MLAT), was developed by Carroll and Sapon (1959) and is based on the measure of four parameters: phonetic coding ability, grammatical sensitivity, rote learning ability, and inductive learning ability. Carroll's aptitude model was later updated in a way that grouped grammatical sensitivity and inductive learning ability under the concept of language-analytic ability (Skehan 1989). This new conceptualization was motivated by the notion that both components vary in their degree of emphasis rather than in quality (Roehr 2007), since grammatical sensitivity and inductive learning ability are thought to play a similar role in "the identification and generalization of linguistic patterns" (Roehr 2007: 176). Empirical studies have shown that performance and classroom learning in L2 correlate positively with the MLAT, with Pearson's ranging between 0.40 to 0.60 (Hummel 2009).

Nowadays, foreign language aptitude researchers and specialists recognize that new learning circumstances have resulted in a need to upgrade the components of aptitude tests. For instance, the current emphasis on a communicative approach to language learning is in stark contrast with the audiolingual classroom tradition (Hummel 2009). Such considerations led Robinson (2002: 215) to suggest the need for "a measure of aptitude that reflects the processing demands of simultaneous attention to form and meaning, with its attendant demands on working memory". The incorporation of concepts specific to the modern day classroom environment and learning situation is illustrated through the recent creation by Grigorenko, Sternberg and Ehrman (2000) of an aptitude test called the Cognitive Ability for Novelty in Acquisition of Language as applied to foreign language test (CANAL-FT). Unlike the MLAT, which was rooted in psychometric test development, the CANAL-FT is theory driven and based on Sternberg's triarchic theory of human intelligence (Dörnyei 2005; Sternberg 1997; Sternberg 2002), also known as the "theory of successful intelligence" because it includes the cognitive skills "that are necessary for success in everyday life" and not only in a school enviroment (Dörnyei 2005: 50). Thus, a main aim of the CANAL-FT is to measure how people deal with novelty and ambiguity (Sáfár/Kormos 2008) during the learning process. It was proposed that this could be achieved "by gradually introducing a simulated language embedded in a multifaceted language context" (Grigorenko et al. 2000: 392) where learners are required to complete a number of short-learning assignments that involve five knowledge acquisition processes (Dörnyei 2005: 51):

1. Selective encoding: Distinguishing between more and less relevant information for one's purposes.

2. Accidental encoding: Encoding background or secondary information and grasping the background context of the information stream.

3. Selective comparison: Determining the relevance of old information for current tasks to enhance learning.

4. Selective transfer: Applying encoded or inferred rules to new contexts and tasks.

5. Selective combination: Synthesizing the disparate pieces of information that have been collected via selective and accidental encoding.

These acquisition processes are operationalized "at the lexical, morphological, semantic and syntactic levels and in two language modes of input and output: visual and oral" (ibid.). Another important element of this test is that it measures working and long-term memory through two types of recall tasks: "immediate recall right after learning takes place" and" delayed recall at some substantial time interval" after learning has occurred (ibid.).

A short version of this test denoted CANAL-S has been selected for this study and will be described in more details in the Methodology section.

2.2   Foreign language anxiety

The notion that affective variables such as anxiety, motivation, empathy and attitude could have a significant influence in language acquisition originated in the field of psychology (Gardner 1960; Lambert 1963), but initially did not find a large following in L2 acquisition research. Later, this idea gained more acceptance through the influential work of Stephen Krashen, whose affective filter hypothesis credited emotional factors with the ability to make the language learning input more or less comprehensible to language learners (Krashen 1981). Against this backdrop, Horwitz, Horwitz and Cope (1986) proposed that a situation-specific anxiety construct, denoted foreign language anxiety, induced worry and negative emotions that could impede language learning. It is in an effort to further investigate this concept that Horwitz and her colleagues designed the Foreign Language Anxiety Scale (FLCAS). Since then, the detrimental effects of foreign language anxiety on language learning have been demonstrated in many studies (Horwitz 1986; Aida 1994; Saito/Samimy 1996; Kitano 2001). Recent research (Gregersen 2005) has also shown that anxiety effects could go beyond language related performance to affecting the whole person, as anxious learners display limited facial activity, maintain less eye contact with the teacher, are more rigid and used fewer gestures.

It is noteworthy that dissenting views have been expressed regarding the nature of the causal relationship between foreign language anxiety and L2 learning. A first position has been that "anxiety can play a significant causal role in creating individual differences in both language learning and communication" (MacIntyre 1995: 90). In contrast, Onwuegbuzie, Bailey and Daley (2002) suggested a reciprocal nature between anxiety and achievement "where anxiety in not only a consequence but also a potential cause of FL performance" (p. 43). At the other end of the spectrum, language anxiety has been seen entirely as a result of poor L1 learning ability (Argaman/Abu-Rabia 2002; Ganschow/Sparks 1996; Sparks/Ganschow 1991, 1995, 2007), as it has been argued that "there may be confounding interactions between an affective variable such as anxiety and students' native [...] language skills" (Sparks/Ganschow 2007: 263). This view has been contested by other researchers (e.g., Horwitz 2000; MacIntyre 1995; Yan/Horwitz 2008), who state for instance that the "pervasiveness of anxiety and its existence in successful language learners disproves the contention that all anxiety stems from learning disabilities" (Yan/Horwitz 2008: 153). At present, the exact role played by anxiety in SLA is still debated.

2.3   Language learning strategies

Language learning strategies are metacognitive factors that can be defined as " [goal-oriented] techniques or procedures that facilitate a learning task" (Chamot 2001: 25) and are present in every phase of the second language acquisition process. These strategies are important to SLA because:

1.   They lead to a better understanding of the cognitive, social, and affective processes in language learning.

2. It raises the possibily that less successful learners could be taught to use strategies in way that emulates their more successful peers (Chamot 2001).

Learning strategy research has investigated the use of different strategies by various types of language learners, as well as their relationships to performance and achievement. For instance, Gan, Humphreys and Hamp-Lyons (2004: 236) claimed that "the more successful students used a wider range of activities than their counterparts". In a more nuanced way, Oxford (1996) suggested that unsuccessful students do not use fewer learning strategies than the more successful ones, but tend instead to use them inadequately, i.e., without taking into account their own learning style preferences or the nature of the task at hand.

Over the years, several general language learning strategy inventories have been proposed (Cohen 1998; O'Malley/Chamot 1990; Oxford 1990). Studies have indicated a connection between Oxford's (1990) Strategy Inventory for Language Learning (SILL) and L2 performance scores (Mullins 1992; Olivares-Cuhat 2002), and it is generally accepted that the use of language learning strategies are linked with achievement in the L2 classroom and help students reach learner independence (Vidal 2002: 44). Investigators have also pointed out that factors such as proficiency (Magogwe/Oliver 2007), age (Magogwe/Oliver 2007), nationality (Griffiths 2003), course level (Griffiths 2003) and gender (Green/Oxford 1995) are among those that affect language learning strategy use and choice among all language learners.

2.4   Relationships between learner variables

Surprisingly, only few attempts have previously been made to analyze the concurrent influence of several learner factors on L2 proficiency and performance. As pointed out by Gardner, Tremblay and Masgoret (1997: 344): "there is a lack of research examining all these [learner] variables simultaneously". In this respect, earlier investigations may be roughly divided into two types of approach:

1. The main focus is to identify functional dependencies that link different types of variables. As a first example, Bialystok and Fröhlich (1978) analyzed the role of four independent variables (aptitude, attitude and motivation, field independence, and learning strategies) with regard to four measures of language achievement (reading, writing, listening, and grammar). These observations led to a relationship model in which aptitude and field independence were related to strategy use and attitude, while aptitude and strategy use affected achievement. In another paper, Gardner et al. (1997: 353) measured a large variety of individual differences and connected them by means of a causal model. As a result, they found "specific correlational and causal paths linking seven latent variables: a) language attitudes; b) language aptitude, c) field independence, d) motivation, e) learning strategies, f) self-confidence, and g) language achievement".

2. The aim is to quantify and rank the relative role played by various learner variables towards proficiency of performance. Along these lines, a study conducted by Ehrman and Oxford (1995) recorded correlation values between individual difference variables and performance as derived from reading and speaking proficiency scores. The results suggested that cognitive aptitude was the strongest correlate of learning success. In a more recent study, Onwuegbuzie et al. (2001) investigated the relationships between 18 learner variables and L2 performance based on a sample of 184 students. As they pointed out the basic limitations of the correlation estimates used by Ehrman and Oxford (1995) to analyze multivariate data, they relied instead on a linear regression analysis to identify these relationships. As a further step, squared semipartial coefficients and squared partial correlation coefficients were estimated "to assess each variable's individual contribution" (Onwuegbuzie et al. 2001: 8). From this approach,

it was concluded that overall academic achievement (a measure of cognitive abilities) was the most significant contributor of performance, closely followed by foreign language anxiety.

3   Methodology

3.1   Participants

Forty students enrolled in beginning Spanish courses took part to this study on a voluntary basis. The sample consisted of 10 males and 30 females whose ages ranged from 18 to 27, with a mean of 21.2 and a standard deviation of 2.0. The data was gathered at a large midwestern university during a summer term. All participants signed an informed consent form and were provided with a monetary compensation for their efforts.

3.2   Performance metrics

All students took the same tests under the supervision of the course coordinator. In addition, a unified grading criteria was established by the department. An average between mid-term and final exam grades (expressed in terms of percentages) was used as measure of performance.

3.3   Learner variables metrics

The CANAL-S, FLCAS and SILL were administered in order to evaluate language learning aptitude, anxiety and learning strategy use. In addition, demographical data was gathered at the beginning of the semester by means of a background questionnaire.

3.3.1   Cognitive ability for novelty in acquisition of language – short version (CANAL-S)

The CANAL-S measures the aptitude to learn an L2 by means of a simulation test in which participants are expected to learn elements of a new artificial language, called Ursulu (Grigorenko et al. 2000). This simulated language gathers several aspects of existing languages, but does not actually sound like any specific one. The value of this measure lies in the satisfactory psychometric properties of its content, construct, and criterion validity (Grigorenko et al. 2000). It is a paper-and-pencil test with visual and auditory stimuli, which takes about two hours to be administered in its short version. The test consists of nine sections (five of which relate to immediate recall and four to delayed recall questions), which are grouped in five parts:

1. Learning meanings of neologisms from context (immediate recall and delayed recall).
2. Understanding the meaning of passages (immediate and delayed recall).
3. Continuous paired-associate learning (immediate and delayed recall).
4. Sentence inference (immediate and delayed recall).
5. Learning language rules (immediate recall).

The responses given during each part and averages evaluated over immediate and delayed recall questions combine to yield 7 scores.

3.3.2   Foreign language classroom anxiety scale (FLCAS)

The FLCAS was chosen because of its high reliability (alpha coefficient = 0.93 and a test-retest reliability over eight weeks of $r = 0.83, p < 0.001$), internal consistency, and widespread research use (Horwitz et al. 1986; Liu 2006; Rodríguez/Abreu 2003; Yan/Horwitz 2008). This

instrument consists of 33 items with responses ranging from strongly agree to strongly disagree on a 5-level Likert scale. It is a paper-and-pencil survey which takes approximately 15 minutes to be completed.

3.3.3   Strategy inventory for language learning (SILL)

The SILL version 5.1 was selected to assess language learning strategy use. The choice of this instrument was motivated by its high reliability (0.93 Kronbach alpha), widespread acceptance (Weaver/Cohen 1997) and friendly paper-and-pencil format. The SILL consists of 80 items that have been grouped into six subscales corresponding to the following learning strategies: memory, cognitive, compensation, metacognitive, affective, and social strategies. For simplicity, the related scores are denoted in Tables by the letters A, B, C, D, E and F, respectively. Answers are given on a 1-to-5 Likert scale format in which strategy use is rated from "almost never true of me" to "almost always true of me". Approximately thirty minutes are required to complete the survey.

3.4   Demographic data

Age and gender of the test subjects were also recorded.

3.5   Data analysis

3.5.1   Causal model

The diagram shown in Figure 1 gives a representation of the causal model used in this study. It identifies learner variables belonging to different categories (such as cognitive, affective, metacognitive and demographic) as potential causes of variations observed in L2 performance, a view consistent with the approach proposed by Ehrman and Oxford (1995) and Onwuegbuzie et al. (2001). Within this framework, the strengths of the links between individual difference variables and L2 performance can be quantified with a linear fit yielding the coefficients β_i (Cohen/Cohen/West/Aiken 2003). Otherwise, the correlation coefficients r_{ij} between learner variables are denoted in the diagram by curved lines. Oftentimes, the learner variables are sufficiently correlated for the r_{ij} to be significantly different from zero, a fact that is at the root of most of the difficulties encountered when identifying the parameters for this type of model:
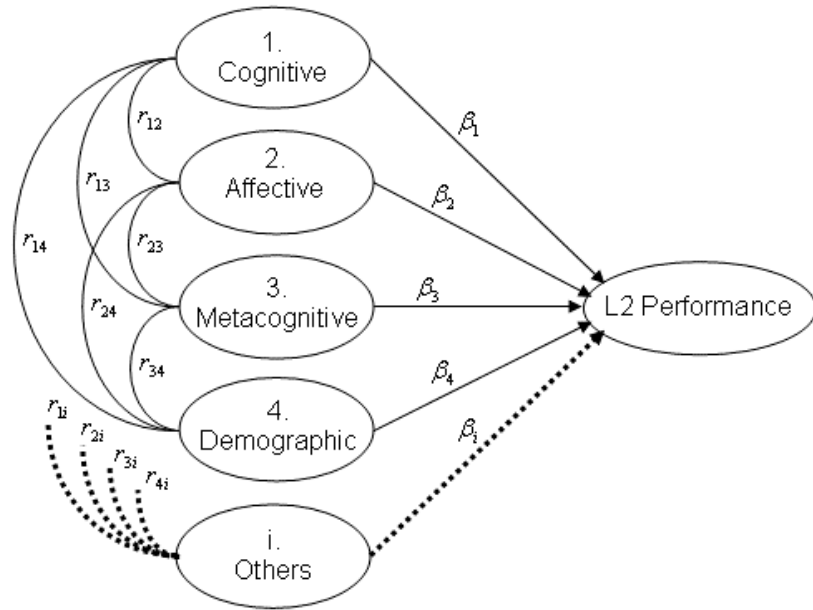
Figure 1: Causal Model

1. If the r_{ij} are different from zero, then the use of a bivariate analysis (such as Pearson product-moment correlation) to evaluate the coefficients β_i would be inaccurate. Instead, it is necessary to use a multiple regression analysis in order to simultanously take the effects of all the variables into account (Cohen et al. 2003).

2. The number of correlated variables used in a linear regression must be limited, as it could otherwise lead to insufficient statistical power (Cohen 1988) and model overfit (Montgomery et al. 2001). An adequate selection of the variables may be achieved by power analysis (as discussed in section 3.6.4), stepwise regression (SPSS 1999) and all-possible-regressions procedure (Montgomery et al. 2001).

3. While the main aim of this paper is to estimate the relative importance of learner variables with respect to L2 performance, the meaning of this idea is not obvious if the causal variables are correlated. The definition of an appropriate metric is discussed in some details in section 3.5.2.

4. The validity of a causal model is subject to the condition that the independent variables are not correlated to an unexpressed variable that may reflect the true causal mechanism (Cohen et al. 2003). An in-depth analysis of this type of confounding effects would go beyond the scope of this paper, but it is possible to evaluate the robustness of the model against this type of interaction (see section 3.5.3).

3.5.2   Estimators for relative importance

The ambiguity associated with the concept of relative importance can be illustrated by the Venn diagram shown in Figure 2, where the variance of each variable is represented by a unit circle. In this picture, $a + b + c + e$ is the total Y variability and $a + b + c$ is the Y variability explained by X_1 and X_2 (Cohen et al. 2003). From this figure it is clear that if X_1 and X_2 are correlated, then their individual contribution to Y (the common area c) is not uniquely defined. The problem is then to approporion the role of each variable in an equitable manner. A first option would be to estimate the percentage increase in Y variability as the variables X_i are added successively into the model. The relative importance of X_1 and X_2 would then amount

to *a+c* and *b*, respectively. However, this outcome is somewhat arbitrary in that it depends on the order of introduction of the variables. A way to solve this issue was proposed by Lindenman, Merenda and Gold (1980) and consists of taking the average of the contributions (or semipartial coefficients) of X_i for all possible orderings, a metric denoted here as LMG. This approach was chosen in this paper instead over other measures of relative importance, such as the Pratt index (Pratt 1987), as it automatically takes into account all the variability explained by the X_i variables and does not yield negative values (Grömping 2006). A statistical software was used in this study due to the complexity of computing the LMG metric and its related confidence intervals (The R Project for Statistical Computing, Version 2.9.2; Relaimpo: Relative Importance of Regressors in Linear Models, Version 2.1–2).
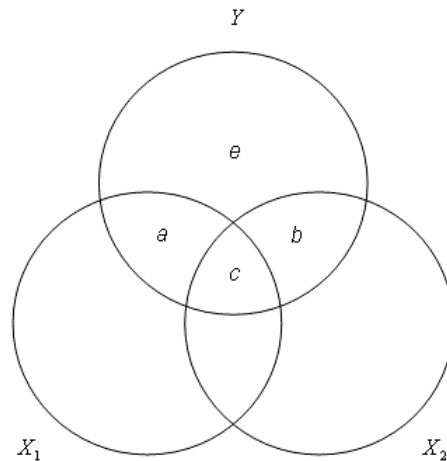

Figure 2: Venn Diagram of Variances in a Multiple Linear Regression

3.5.3   Impact of an unmeasured confounding variable

A potential issue with the model shown in Figure 1 is that the regression coefficents β_i found between learner factors and L2 performance may not be associated with causal mechanisms, but mere correlations with an unmeasured confounding variable (CV). Normally, the identification of such cause-and-effect relationships are best addressed within the context of an experimental study (Winer et al. 1991). However, the techniques required by this approach (randomization, statistical control of all potential confounds) are not easily applicable to research pertaining to individual differences in L2 acquisition. For instance, it would be unethical to submit students to a range of various treatments for the sake of randomization. Another limitation is that the prospect of controlling every potential confounding variable (CV) is not realistic in a field where the background of the subjects is likely to be very diverse. However, it is possible to assess the robustness of a causal inference with respect to the inclusion of an unmeasured CV within the framework of an observational study with a statistical method developed by Frank (2000). In this approach, a minimum impact threshold above which the validity of an inference is altered by a CV is calculated in function of the sample size, number of parameters included in the model, significance level and correlation coefficient found between dependent and independent variables. This threshold can be expres-sed in the form of a set of values required for the correlation coefficients between CV and independent variable on the one hand, and between CV and dependent variable on the other hand (Calculating indices, Version 3). It should be emphasised that the goal of this procedure is not to prove causality, but rather to quantify the risk that an unmeasured variable may invalidate this assumption.

3.6   Power analysis

Given the relatively small number of participants (N = 40), a power analysis was in order to determine whether the sample size was sufficient for the purpose of this pilot study. To this end, typical effect sizes were derived from results previously published in this field of research. Powers were then estimated for Pearson product-moment correlation coefficients and regression models. A significance level of 5% and non-directional behaviors (two-tailed tests) were assumed in all cases. Computations were carried out with a statistical software (The R Project for Statistical Computing, Version 2.9.2) using the formulas established by Cohen (1988).

3.6.1   Correlation between CANAL-S and L2 performance

The author is not aware of large scale investigations in which Pearson's correlations were evaluated between a version of the CANAL test and student's performance. However, a study aimed at assessing the external construct validity of the CANAL-FT indicated that the subscales of this instrument were highly correlated with the MLAT (Grigorenko et al. 2000). Otherwise, MLAT correlations ranging from 0.52 to 0.54 with respect to learning success were found by Ehrman (1998) for Western languages. Consequently, it seems reasonable to assume an effect size above 0.40 for the CANAL-S, thus implying a power greater than 73% for this test in our study.

3.6.2   Correlation between FLCAS and L2 performance

Pearson's r between FLCAS and students' performance ranging from -0.41 to -0.47 were reported by Saito and Samimy (1996) for intermediate and advanced students of Japanese. Otherwise, Horwitz (1986) measured r values between FLCAS and final grades of $r = -0.49$ in two beginning Spanish classes, and $r = -0.54$ in two French classes. Transposed to this study, these effect sizes correspond to powers ranging from 76% to 96%.

3.6.3   Correlation between SILL and L2 performance

In a study conducted with the SILL by Mullins (1992) involving 110 EFL majors at a Thai university, it was reported that L2 performance was significantly correlated with compensation strategy use at $r = 0.32$ and with metacognitive strategy use at $r = 0.24$. For the present study, these effect sizes translate into powers amounting to 52% and 32%, respectively.

3.6.4   Multiple regression models for the prediction of L2 performance

In Onwuegbuzie et al. (2001), a regression analysis was conducted using various cognitive, affective, personality and demographic variables to predict L2 performance. The data was obtained from a sample of 184 students enrolled in Spanish, French, German and Japanese. As a result, a model explaining 32.4% of the course grades variability was derived, including two cognitive variables (academic achievement and students' expected level), one affective variable (foreign language anxiety), one personality variable (cooperativeness) and one demographic variable (gender). Given the smaller sample size of this study and in order to avoid overfitting, the aim is to build a model involving only three variables. In order to estimate the corresponding effect size of a model only including a cognitive, affective and demographic variable, an equivalent overall R^2 value was calculated with the following formula (Cohen et al. 2003: 83):

$$R^2 = \sum_{k=1}^{3} \beta_k r_{Yk} \, ,$$

where the β_k and r_{Yk} stand for the standardized regression coefficients and correlation coefficients between independent and dependent variables reported in Onwuegbuzie et al.'s (2001). As a result, it was found that R^2 amounted to 0.241, so that with a sample size of 40, a statistical power of 82% was derived.

3.6.5 Summary

Overall, the power analysis showed that, in spite of a relatively small sample size (N = 40), the estimated powers (i.e., the probability of avoiding type II errors) were close to or exceeded 80% for statistical tests to be conducted on L2 aptitude and anxiety with respect to L2 performance. The strong effect sizes associated with these variables explained the high power levels that were obtained in spite of a relatively small sample size. In the case of language learning strategies, the highest power amounted to only 52% in the case of compensation strategies. Given the relative simplicity of administering the SILL, it was still decided to include this test in the study even though the probability of avoiding Type II errors is moderate in this case.

4   Results

First, the normality of the distributions was tested with Kolmogorov-Smirnov test (SPSS 1999). It was found that this assumption was verified for all the variables, thus justifying the use of the statistical methods described in the previous section.

Pearson product-moment correlation coefficients were then evaluated between the measures provided by the instruments. First, it was observed that the subscales within CANAL-S and SILL were strongly correlated (see Tables 1 and 2). This behavior is expected from these instruments (Grigorenko et al. 2000; Oxford/Ehrman 1995) and indicates that distinctions made by the various subscales are not so clear-cut. Next, it was found that various learner variables across categories are also significantly correlated (see Table 3): The FLCAS scores correlated negatively with SILL B (cognitive subscale) (r = -0.406, p < 0.05), SILL D (metacognitive) (r = -0.377, p < 0.05), as well as negatively again with the CANAL-S section 5 (learning language rules) (r = -0.335, p < 0.05) and CANAL-S section 4 (sentence inference) (r = -0.345, p < 0.05). Moreover, SILL A (memory) correlated with CANAL-S section 4 (sentence inference) (r = 0.380, p < 0.05) and section 5 (r = 0.342, p < 0.05), and SILL B (cognitive) correlated with CANAL section 4 (sentence inference) (r = 0.448, p < 0.05). Otherwise, some learner variables correlated significantly with L2 performance. Indeed, grade average correlated with CANAL-S section 1 (learning meanings of neologisms) (r = 0.324, p < 0.05), section 4 (sentence inference) (r = 0.388, p < 0.05), section 5 (learning language rules) (r = 0.369, p < 0.05), immediate recall (r = 0.461, p < 0.05), and FLCAS (r = -0.645, p < 0.01) – see Table 4.

Subscale	SILL A	SILL B	SILL C	SILL D	SILL E	SILL F
SILL A						
SILL B	0.393*					
SILL C	0.326*	0.272				
SILL D	0.396*	0.684**	0.404*			
SILL E	0.304	0.044	0.140	0.380*		
SILL F	0.444**	0.321*	0.141	0.399*	0.463**	
Overall	0.699**	0.799**	0.468**	0.853**	0.453**	0.645**

Table 1: Pearson's *r* between SILL subscales
*p < 0.05, **p < 0.01

Subscale	Section 1	Section 2	Section 3	Section 4	Section 5	Immediate Recall
Section 1						
Section 2	0.149					
Section 3	0.304	0.299				
Section 4	0.418**	0.148	0.429**			
Section 5	0.229	0.158	0.261	0.560**		
Immediate Recall	0.645**	0.295	0.655**	0.751**	0.706**	
Delayed Recall	0.382*	0.621**	0.518**	0.671**	0.544**	0.566**

Table 2: Pearson's *r* between CANAL-S subscales
*p < 0.05, **p < 0.01

Learner Variable	CANAL-S Section 4	CANAL-S Section 5	CANAL-S Immediate Recall	FLCAS
FLCAS		-0.335*	-0.345*	
SILL A	0.380*	0.342*		
SILL B	0.448*			-0.406*
SILL D				-0.377*

Table 3: Significant Pearson's *r* between different types of learner variables
*p < 0.05, **p < 0.01

Learner Variable	CANAL-S section 1	CANAL-S section 4	CANAL-S section 5	CANAL-S Immediate Recall	FLCAS
SGPA		0.399*		0.444*	-0.626**
Grade average	0.324*	0.388*	0.369*	0.461*	-0.645**

Table 4: Significant Pearson's *r* between learner variables and measures of L2 performance
*p < 0.05, **p < 0.01

As pointed out in section 3.5.1, the correlations found between various learner variables make it necessary to use a multiple regression analysis to evaluate the parameters of the causal model. With a stepwise procedure, a first linear regression was obtained in which FLCAS, CANAL-S immediate recall and SILL F (social) predicted grade average with an adjusted R^2 amounting to 48.5% – see Table 5. In addition, an all-possible-regressions procedure showed that a second linear model of grade average could be derived in function of FLCAS, CANAL-S immediate recall and age with an adjusted R^2 of 48.4% – see Table 6. It should be noted that stepwise procedure and power analysis (see section 3.6.4) both suggested the use of three independent variables for this study.

Model	Unstandardized Coefficients *B*	Standard Error	Standardized Coefficients β	*t* Value
Constant	98.823	13.859		7.131***
FLCAS	-0.243	0.056	-0.535	-4.354***
CANAL-S immediate recall	0.702	0.270	0.320	2.606*
SILL F	-4.940	2.202	-0.260	-2.243*

Table 5: First regression model obtained with grade average as dependent variable
R^2 = 0.524, F(3,36) = 13.233; Adjusted R^2 = 0.485; *p < 0.05, ***p < 0.001

Model	Unstandardized Coefficients *B*	Standard Error	Standardized Coefficients β	*t* Value
Constant	120.518	20.611		5.847***
FLCAS	-0.240	0.056	-0.530	-4.310***
CANAL-S immediate recall	0.644	0.269	0.293	2.390*
Age	-1.689	0.758	-0.257	-2.229*

Table 6: Second regression model obtained with grade average as dependent variable
R^2 = 0.524, F(3,36) = 13.195; Adjusted R^2 = 0.484; *p < 0.05, ***p < 0.001

LMG metrics and related confidence intervals were then evaluated from the multiple regression models – see Tables 7 and 8. Confidence intervals (CI) of 80% and 90% were selected for the purpose of this pilot study.

Learner Variable	Relative Importance	Lower of 80% CI	Upper of 80% CI	Lower of 90% CI	Upper of 90% CI
FLCAS	67.1%	49.8%	82.9%	42.7%	85.8%
CANAL-S immediate recall	26.3%	15.1%	41.3%	12.2%	46.0%
SILL F	6.6%	1.2%	20.4%	0.007%	26.3%

Table 7: LMG measures of relative importance derived from first regression model

Learner Variable	Relative Importance	Lower of 80% CI	Upper of 80% CI	Lower of 90% CI	Upper of 90% CI
FLCAS	66.2%	47.4%	80.5%	40.8%	83.2%
CANAL-S immediate recall	24.9%	14.4%	40.9%	11.7%	46.2%
Age	8.9%	2.4%	21.1%	1.5%	25.3%

Table 8: LMG measures of relative importance derived from second regression model

According to the results shown in Tables 7 and 8, the measure of foreign language anxiety is a prominent predictor of performance. The task of the researcher is then to examine the likelihood of alternate causal factors, i.e., to assess the potential effect of an unmeasured confounding variable. To this end, minimal correlations above which foreign language anxiety would be disqualified as a causal factor on L2 performance were evaluated with Frank's approach (2000) and are shown in Table 9.

Model Number	Minimum Correlation Between Potential Confounder and Measure of Language Learning Anxiety	Minimum Correlation Between Potential Confounder and Measure of Performance
1	0.59	0.65
2	0.59	0.65

Table 9: Minimal regression coefficients required to invalidate causal inference of FLCAS on performance

5   Discussion

The primary goal of this pilot study was to estimate the relative importance of some significant learner variables with respect to L2 performance. A second aim was to weigh in the possibility that a confounding variable invalidated the assumption of a causal link between foreign language anxiety and L2 performance.

As general comments, it should be first noted that cognitive, affective and metacognitive variables were sufficient to explain 48.5% of the variability in exam grade average (see Table 5). This is consistent with the suggestion made by Horwitz (2008) that these three variable catgories are especially relevant to the L2 teacher. Second, the presence of the immediate recall component of the CANAL-S in the model illustrated the importance of working memory to the language learning process. Third, while the signs of the regression coefficients associated with language aptitude (positive) and anxiety (negative) were expected, the negative coefficient linking the use of social strategy with L2 performance is counterintuitive.

However, this result is not inconsistent with Oxford's warning that the positive effect of learning strategies depends less on their frequency than on their adequate use (1996).

As to the main question of ranking the relative importance of learner variables with respect to L2 performance, the results given in Tables 7 and 8 with 90% CI indicate that the outcome of the comparison between language learning aptitude and anxiety was inconclusive since the CI overlapped (albeit only slightly). By the same token, language learning aptitude was not found to be more important than social strategy use (SILL F) in the first model and age in the second model. However, it was shown that foreign language anxiety was more important with respect to L2 performance than the learning language strategies and demographic variables included in this study. When lowering the CI requirement to 80%, foreign language anxiety was then found to be a more important contributor than aptitude. Several remarks follow from these findings:

1. The CI turned out to be relatively large, leaving unanswered the question of ranking the relative importance in cases where differences seemed at first to be quite significant. One reason for this outcome could be that the use by the statistical sofware of a bootstrap algorithm to calculate the CIs could have led to overly wide estimates (Grömping 2006). However, the sample size also played a role in this respect and this pilot study seems to indicate that investigations of relative importance could require high numbers of participants to yield significant results.

2. Language learning anxiety did stand out as the most influential factor within the framework of this study. This conclusion was somewhat unexpected, as aptitude is often viewed as the most effective language learning variable. Possibly, this discrepancy could be explained by a lower effectiveness of the CANAL-S to measure aptitude than with more traditional tests such as the MLAT. As another possible reason, the specific conditions found in this classroom setting may have led to a higher sensitivity of L2 perfomance to language learning anxiety than to aptitude.

3. Even though several factors were not retained as predictors and therefore were not assigned any relative importance, it does not necessarily imply that these variables were unimportant. For instance, the fact that gender was not ranked in this study is consistent with Ehrman and Oxford (1995), who had found no correlation between this parameter and foreign language proficiency, but it deviates from the findings of Onwuegbuzie et al. (2001). Thus, it would appear that more research is needed to clarify the role of gender on L2 performance. Otherwise, even though only one subscale of the CANAL-S and one subscale of the SILL were included in the linear regressions, it was shown that a large amount of redundancy existed within these instruments (high correlation coefficients were reported in Tables 1 and 2), so that some statistical variations could easily have led to the selection of different subscales to represent the contributions of language learning aptitude and strategy use to L2 performance.

As to the assumption of a cause-and-effect relationship between foreign language anxiety and L2 performance, Frank's method (2000) shows that an unmeasured variable could have had a confounding effect provided that its correlation coefficients with foreign language anxiety and L2 performance exceeded the set of values given in Table 17. To ascertain the likelihood of such a possibility, one can take into account the suggestion that such a confounder is a type of cognitive variable, since it would be tied to previous L1 achievements (Sparks/Ganshow 1991, 1995, 2007). In such a case, it would then seem plausible that the correlation coefficient of an unmeasured cognitive variable with L2 performance exceeded the value of 0.65 given on the right-hand column of Table 17, since one can expect to find strong links between the measure of cognitive skills and language learning success (Carroll 1990; Ehrman 1998).

However, it appears much less likely that such a cognitive variable would be highly correlated with the FLCAS (the measure of an affective variable), as would be required by the left-hand column of Table 17. Therefore, some doubt was cast on the idea that a confounder invalidated the assumption of a causal link between language learning anxiety and L2 performance.

6   Implications for future research

In this pilot study, language learner variables were ranked as a mean of assessing their importance towards L2 performance. Compared to previous research, a main feature of this method is that it provides a multivariate perspective of the roles played by diverse learner factors.

As a first implication, the central and somewhat dual roles of cognitive and affective variables was confirmed, suggesting that both aspects should be taken into consideration when assessing the effects of individual differences on L2 performance.

Second, this approach allows the researcher to relax the requirement that learner variable tests and surveys be designed to only measure a single type of variable (e.g., cognitive, affective, metacognitive and demographic), thus opening the door to the use of more sophisticated tools. A case in point is the CANAL. While the aim of this test is to evaluate language learning aptitude, it is likely for this intrument to spur strong affective reactions as the participants are being exposed to the unfamiliar sounds of an artificial language. At first, this feature may be viewed as a drawback, since the researcher may seek to insulate subjects from factors that are not purely cognitive. However, the nature of the CANAL also enables it to better estimate the role played by working memory, which is a critical element in L2 learning. In this context, the use of the statitical tools featured in this paper was then helpful in separating the overlapping cognitive and affective influences both measured by the CANAL and FLCAS.

Third, the relatively large widths derived for the confidence intervals of the relative importance metrics point to the need to secure large sample size and implement reliable instruments for this type of studies.

References

Aida, Yukie (1994): "Examination of Horwitz, Horwitz and Cope's Construct of Foreign Language Anxiety. The Case of Students of Japanese". *The Modern Language Journal* 78: 155–168.

Argaman, Osnat/Abu-Rabia, Salim (2002): "The Influence of Language Anxiety on English Reading and Writing Tasks among Native Hebrew Speakers". *Language, Culture and Curriculum* 15*:* 143–160.

Bialystok, Ellen/Fröhlich, Maria (1978): "Variables of Classroom Achievement in Second Language Learning". *The Modern Language Journal* 62/7: 327–335.

Calculating Indices (Version 3) [Computer software]. Available at https://www.msn.edu/~kenfrank/papers/calculating%20indices%203.xls (accessed February 10, 2010).

Carroll, John/Sapon, Stanley (1959): *Modern Language Aptitude Test (MLAT) Manual*. San Antonio, TX.

Chamot, Anna (2001): *Learner Contributions to Language Learning*. New York, NY.

Cohen, Andrew (1998): *Strategies in Learning and Using a Second Language*. London, UK.

Cohen, Jacob (1988): *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ.

Cohen, Jacob et al. (2003): *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Mahwah, NJ.

Dornyei, Zoltan (2005): *The Psychology of the Language Learner*. Mahwah, NJ.

Ehrman, Madeleine (1998): "The Modern Language Aptitude Test for Predicting Learning Success and Advising Students". *Applied Language Learning* 9*:* 31–70.

Ehrman, Madeleine/Oxford, Rebecca (1995): "Cognition Plus. Correlates of Language Learning Success". *The Modern Language Journal 79:* 67–89.

Frank, Ken (2000): "Impact of a Confounding Variable on a Regression Coefficient". *Sociological Methods and Research* 29*:* 147–194.

Gan, Zhengdong et al. (2004): "Understanding Successful and Unsuccessful EFL Students in Chinese Universities". *The Modern Language Journal* 88/2: 229–244.

Ganshow, Leonore/Sparks, Richard (1996): "Anxiety about Foreign Language Learning among High School Women". *The Modern Language Journal* 80/2: 199–212.

Gardner, Robert (1960): *Motivational Variables in Second Language Acquisition.* Unpublished doctoral disseration. McGill University, Canada.

Gardner, Robert et al. (1997): "Towards a Full Model of Second Language Learning: an Empirical Investigation". *The Modern Language Journal* 81/3: 344–362.

Green, John/Oxford, Rebecca (1995): " A Closer Look at Learning Strategies, L2 Proficiency and Gender". *TESOL Quarterly* 29/2: 261–297.

Gregersen, Tammy (2005): "Non-Verbal Cues: Clues the Detection of Foreign Language Anxiety". *Foreign Language Annals* 38: 388–397.

Griffiths, Carol (2003): "Patterns of Language Learning Strategy Use". *System* 31/3: 367–383.

Grigorenko, Elena et al. (2000): "A Theory-Based Approach to the Measurement of Foreign Language Learning Ability. The CANAL-F Theory and Test". *The Modern Language Journal* 84/4: 390–405.

Grömping, Ulrike (2006): " Relative Importance for Linear Regression in R. The Package Relaimpo". *Journal of Statistical Software* 17: 1–26. Available at http://www.jstatsoft.org/ v17/i01 (accessed February 20, 2010).

Horwitz, Elaine (1986): "Preliminary Evidence for the Reliability and Validity of a Foreign Language Anxiety Scale". *TESOL Quarterly 20*: 559–562.

Horwitz, Elaine (2000): "Horwitz Comments. It ain't Over 'til it's Over. On Foreign Language Anxiety, First Language Deficits, and the Confounding of Variables". *The Modern Language Journal* 84/2: 256–259.

Horwitz, Elaine (2001): "Language Anxiety and Achievement". *Annual Review of Applied Linguistics* 21: 112–126.

Horwitz, Elaine (2008). *Becoming a Language Teacher*. Boston, MA.

Horwitz, et al. (1986): "Foreign Language Classroom Anxiety". *The Modern Language Journal* 70: 125–132.

Hummel, Kirsten (2009): "Aptitude, Phonological Memory, and Second Language Proficiency in Nonnovice Adult Learners". *Applied Psycholinguistics* 30/02: 225–249.

Kitano, Kazu (2001): " Anxiety in the College Japanese Language Classroom". *The Modern Language Journal* 85/4: 549–566.

Krashen, Stephen (1981): *Principles and Practice in Second Language Acquisition. English Language Teaching Series*. London.

Lambert, Wallace (1963): "Psychological Approaches to the Study of Language. Part I: On Learning, Thinking and Human Abilities". *The Modern Language Journal* 3: 51–66.

Lindeman, Richard et al. (1980): *Introduction to Bivariate and Multivariate Analysis.* Glenview, IL.

Liu, Meihua. (2006): "Anxiety in Chinese EFL Students at Different Proficiency Levels". *System: An International Journal of Educational Technology and Applied Linguistics* 34: 301–304.

MacIntyre, Peter (1995): "How does Anxiety Affect Second Language Learning? A Reply to Sparks and Ganschow". *The Modern Language Journal* 79: 90–99.

MacIntyre, Peter (1999): "Language Anxiety. A Review of the Research for Language Teachers". In: Young, Dolly J. (ed.): *Affect in Foreign and Second Language Learning*. Boston: 24–45.

Magogwe, Joel/Oliver, Rhonda (2007): "The Relationship between Language Learning Strategies, Age and Self-Efficacy Beliefs. A Study of Language Learners in Botswana". *System. An International Journal of Educational Technology and Applied Linguistics* 35: 338–352.

Montgomery, Douglas et al. (2001): *Introduction to Linear Regression Analysis*. New York.

Mullins, P. Yampek (1992): *Successful English Language Learning Strategies of Students Enrolled in the Faculty of Arts, Chulalongkorn University, Thailand*. Unpublished doctoral dissertation. United States International University, San Diego.

Olivares-Cuhat, Gabriela (2002): "Learning Strategies, Writing Textbooks and Achievement in the Spanish Classroom: A Case-Study". *Foreign Language Annals* 35/5: 561–570.

O'Malley, John/Chamot, Anna (1990): *Learning Strategies in Second Language Acquisition*. Cambridge, UK.

Onwuegbuzie, Anthony et al. (2001): "Cognitive, Affective, Personality and Demographic Predictors of Foreign-Language Achievement". *The Journal of Educational Research* 94: 3–15.

Onwuegbuzie, Anthony et al. (2002): "The Role of Foreign Language Anxiety and Students' Expectations in Foreign Language Learning". *Research in the Schools* 9: 33–50.

Oxford, Rebecca (1990): *Language Learning Strategies*. Boston.

Oxford, Rebecca (1996): *Language Learning Strategies around the World: Cross-Cultural Perspectives*. Honolulu, HI.

Oxford, Rebecca/Burry-Stock, Judith (1995): "Assessing the Use of Language Learning Strategies Worldwide with the ESL/EFL Version of the Strategy Inventory for Language Learning (SILL)". *System* 23/1: 1–23.

Oxford, Rebecca/Ehrman, Madeleine (1995): "Adults' Language Learning Strategies in an Intensive Foreign Language Program in the United States". *System* 23/3: 359–386.

Pratt, John (1987): "Dividing the Indivisible: Using Simple Symmetry to Partition Variance Explained". In: *Proceedings of Second Tampere Conference in Statistics*. University of Tampere, Finland: 245–260.

Relaimpo: Relative Importance of Regressors in Linear Models (Version 2.1-2) [Computer software]. Available at http://prof.beuth-hochschule.de/groemping/relaimpo/ (accessed February 20, 2010).

Robinson, Peter (2005): "Aptitude and Second Language Acquisition". *Annual Review of Applied Linguistics* 25: 46–73.

Rodríguez, Máximo/Abreu, Orángel (2003): "The Stability of General Foreign Language Classroom Anxiety across English and French". *The Modern Language Journal* 87/3: 365–374.

Roehr, Karen (2007): "Metalinguistic Knowledge and Language Ability in University-Level L2 Learners". *Applied Linguistics* 29/2: 173–199.

Sáfár, Anna/Kormos, Judit (2008): "Revisiting Problems with Foreign Language Aptitude". *International Review of Applied Linguistics* 46: 113–136.

Saito, Yoshiko/Samimy, Keiko (1996): "Foreign Language Anxiety and Language Performance. A Study of Learner Anxiety in Beginning, Intermediate, and Advanced-Level College Students of Japanese". *Foreign Language Annals* 29/2: 239–251.

Skehan, Peter (1989): *Individual Differences in Second-Language Learning*. London.

Sparks, Richard/Ganschow, Leonore (1991): "Foreign Language Learning Difficulties. Affective or Native Language Aptitude Differences?" *The Modern Language Journal* 75: 3–16.

Sparks, Richard/Ganschow, Leonore (1995): "A Strong Inference Approach to Causal Factors in Foreign Language Learning. A Response to MacIntyre". *The Modern Language Journal* 79: 235–244.

Sparks, Richard/Ganschow, Leonore (2007): "Is the Foreign Language Classroom Anxiety Scale Measuring Anxiety or Language Skills?" *Foreign Language Annals* 40/2: 260–287.

SPSS (1999): *SPSS Application Guide*. Chicago.

Sternberg, Robert (1997): *Successful Intelligence*. New York.

Sternberg, Robert (2002): "The Theory of Successful Intelligence and its Implications for Language Aptitude Testing". In: Robinson, Peter (ed.): *Individual Differences and Instructed Language Learning*. Amsterdam: 13–44.

The R Project for Statistical Computing (Version 2.9.2) [Computer software]. Available at http://www.r-project.org/ (accessed February 20, 2010).

Vidal, Rejane (2002): "Is There a Correlation between Reported Language Learning Strategy Use, Actual Strategy Use and Achievement? *Linguagem and Ensino* 5/1: 43–73.

Weaver, Susan/Cohen, Andrew (1997): *Strategies-Based Instruction: A Teacher-Training Manual*. University of Minnesota: Center for Advanced Research on Language Acquisition.

Winer, Benjamin et al. (1991): *Statistical Principles in Experimental Design*. Boston.

Yan, Jackie/Horwitz, Elaine (2008): "Learners' Perceptions of how Anxiety Interacts with Personal and Instructional Factors to Influence their Achievement in English. A Qualitative Analysis of EFL Learners in China". *Language Learning* 58/1: 151–183.

Young, Dolly (1999): *Affect in Foreign Language and Second Language Learning*. Boston.